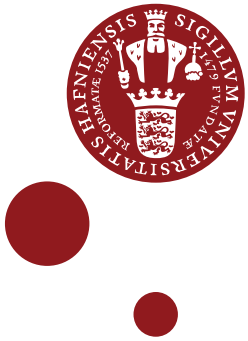
UNIVERSITY OF COPENHAGEN
FACULTY OF HEALTH AND MEDICAL SCIENCES

STATENS
SERUM
INSTITUT



DMSc Thesis

Aspects of genetic diversity, host specificity, and public health significance of single-celled intestinal parasites commonly observed in humans and mostly referred to as ‘non-pathogenic’

Christen Rune Stensvold

Statens Serum Institut
&
Faculty of Health and Medical Sciences
University of Copenhagen

2025

DMSc Thesis

Aspects of genetic diversity, host specificity, and public health significance of single-celled intestinal parasites commonly observed in humans and mostly referred to as 'non-pathogenic'

Christen Rune Stensvold
BMedSc, MSc, PhD

Laboratory of Parasitology
Department of Bacteria, Parasites and Fungi
Infectious Disease Preparedness
Statens Serum Institut
Copenhagen
Denmark

Faculty of Health and Medical Sciences
University of Copenhagen

ISBN: 978-87-7629-109-9

Assessment Committee

Professor Kurt Hanevik, University of Bergen, Norway

Professor David Modrý, Masaryk University, Czech Republic

Professor Birgitte J. Vennervald, University of Copenhagen

Chair of the defence

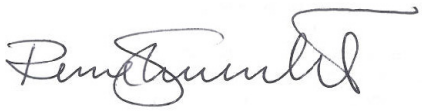
Professor Stig Milan Thamsborg, University of Copenhagen.

The Faculty of Health and Medical Sciences at the University of Copenhagen has accepted this dissertation for public defence for the doctoral degree in medicine. Copenhagen, 6 December 2024.

The defence ceremony will take place Thursday 3 April, 2025, at 13.00 in Foredragssalen, Statens Serum Institut, Copenhagen.

I have wanted to do a compilation of my research for a long time. The process of delimiting the overall topic, however, became a bit lengthy, since my many interests – and the many organisms – have prompted me to try to cover various different aspects of gut parasites during my time as a post doc student and senior scientist based at London School of Hygiene and Tropical Medicine (LSHTM) and Statens Serum Institut (SSI), respectively. This situation may also partly explain the long title, which reflects an attempt to ensure a coherent narrative while still displaying the diversity in my work, focussing on seminal/basic research aspects within gut parasitology rather than the mere development and evaluation of diagnostic assays for detection and differentiation of intestinal parasites. Still, method development is a condition for carrying out studies of seminal character. The number of included articles may appear high; however, most of them are quite short, and some are similar in structure and content.

I'm thankful to the Danish Medical Research Council (Grant ID: 271-09-0251 Danish Agency for Science Technology and Innovation), Marie Skłodowska-Curie Actions (Call: FP7-PEOPLE-2012-CIG; Grant ID: 321614), and the European Union's Horizon 2020 Research and Innovation Programme; Grant ID: 773830; One Health European Joint Programme [PARADISE Consortium]) for financial support. I thank my employers at SSI for allowing me to help guiding research and development activities and my SSI colleagues for securing a good working environment. Thanks to all co-authors and to colleagues in clinical microbiology departments across the country, University of Copenhagen, Technical University of Denmark and LSHTM. I am thankful to have had the opportunity to (co-)supervise many BSc, MSc, MDs, PhD students and EUPHEM fellows – this has been a constant source of energy for me, and it has helped me develop and expand the research topic. I am deeply thankful to my colleagues in the Parasitology Laboratory. Lee O'Brien Andersen is thanked for all the energy that you put into the development of the metabarcoding assay. Lis Lykke Wassmann and Derakhsandeh Seid Moradi receive special thanks for excellent technical assistance and support. Finally, I wish to extend my very special thanks to Henrik Vedel Nielsen, Jaco Verweij, Marianne Lebbad and Graham Clark for continued support and massive inspiration over the years; your scientific approaches and curiosity have had a large influence on my motivation and professional development.



C. Rune Stensvold – Copenhagen, February 3, 2023.

Contents

PREFACE ... 2
LIST OF ABBREVIATIONS ... 5
LIST OF ARTICLES ... 6
LIST OF TABLES ... 8
LIST OF FIGURES ... 10
SUMMARY IN DANISH ... 13
SUMMARY IN ENGLISH ... 16
1. INTRODUCTION ... 19
1.1. Common Luminal Intestinal Parasitic Protists (CLIPPs) – Overall Classification ... 19
1.2. Detection and Differentiation of CLIPPs ... 20
1.3. Cryptic Genetic Diversity ... 22
1.4. Molecular Characterisation of Parasites ... 23
1.5. Gut Microbiota Profiling ... 25
1.6. Aim of Studies ... 26
2. GENETIC DIVERSITY & HOST SPECIFICITY OF CLIPPs OBSERVED IN HUMANS ... 27
2.1. BLASTOCYSTIS ... 30
2.2. DIENTAMOEBA ... 42
2.3. ENTAMOEBA ... 47
2.3.1. Entamoeba coli ... 52
2.3.2. Entamoeba dispar ... 57
2.3.3. Entamoeba gingivalis ... 60
2.3.4. Entamoeba hartmanni ... 63
2.3.5. Entamoeba moshkovskii ... 67
2.3.6. Entamoeba polecki ... 69
2.3.7. Entamoeba in ungulates ... 73
2.4. ENDOLIMAX ... 74
2.5. IODAMOEBA ... 77
3. SUMMARY OF CLIPPs RARELY REPORTED IN HUMANS ... 80
4. CLIPPs AS MARKERS OF GASTROINTESTINAL HEALTH? ... 82
5. DISCUSSION & FUTURE PERSPECTIVES ... 85
6. REFERENCES ... 92
7. TABLES ... 110

LIST OF ABBREVIATIONS

BLAST – Basic Local Alignment Search Tool

Bp – base pairs

CL – conditional lineage

CLIPPs – common luminal intestinal parasitic protists

CML – clinical microbiology laboratory

DNA – deoxyribonucleic acid

FECT – formol ethyl acetate concentration technique

FMT – faecal microbiota transplantation

GI – gastrointestinal

HIV – human immunodeficiency virus

IBD – inflammatory bowel disease

ITS – Internal Transcribed Spacer

kb – kilobases

mb – megabases

MB – mebendazole

MLO – mitochondrion-like organelle

MZ – metronidazole

NA – not available/applicable

NCBI – National Center for Biotechnology Information

NGS – next-generation sequencing

NHP – non-human primate(s)

Nt – nucleotide

OTU – operational taxonomic unit

PCR – polymerase chain reaction

PM – paromomycin

rDNA – ribosomal DNA

RFLP – restriction fragment length polymorphism

RNA – ribonucleic acid

rRNA – ribosomal RNA

SSI – Statens Serum Institut

SSU – small subunit

ST – subtype

SXT – sulfametaxazole

TMP – trimethoprim

WGS – whole genome sequencing

LIST OF ARTICLES

1. **Stensvold CR,** Alfellani MA, Nørskov-Lauritsen S, Prip K, Victory EL, Maddox C, Nielsen HV, Clark CG. Subtype distribution of *Blastocystis* isolates from synanthropic and zoo animals and identification of a new subtype. Int J Parasitol. 2009 Mar;39(4):473-9. doi: 10.1016/j.ijpara.2008.07.006
2. **Stensvold CR,** Christiansen DB, Olsen KE, Nielsen HV. *Blastocystis* sp. subtype 4 is common in Danish *Blastocystis*-positive patients presenting with acute diarrhea. Am J Trop Med Hyg. 2011 Jun;84(6):883-5. doi: 10.4269/ajtmh.2011.11-0005
3. **Stensvold CR,** Alfellani M, Clark CG. Levels of genetic diversity vary dramatically between *Blastocystis* subtypes. Infect Genet Evol. 2012 Mar;12(2):263-73. doi: 10.1016/j.meegid.2011.11.002
4. **Stensvold CR.** Comparison of sequencing (barcode region) and sequence-tagged-site PCR for *Blastocystis* subtyping. J Clin Microbiol. 2013 Jan;51(1):190-4. doi: 10.1128/JCM.02541-12
5. Poulsen CS, Efunshile AM, Nelson JA, **Stensvold CR.** Epidemiological aspects of *Blastocystis* colonization in children in Ilero, Nigeria. Am J Trop Med Hyg. 2016 Jul 6;95(1):175-9. doi: 10.4269/ajtmh.16-0074
6. **Stensvold CR,** Sørland BA, Berg RPKD, Andersen LO, van der Giezen M, Bowtell JL, El-Badry AA, Belkessa S, Kurt Ö, Nielsen HV. Stool microbiota diversity analysis of *Blastocystis*-positive and *Blastocystis*-negative individuals. Microorganisms. 2022 Jan 31;10(2):326. doi: 10.3390/microorganisms10020326
7. **Stensvold CR,** Arendrup MC, Mølbak K, Nielsen HV. The prevalence of *Dientamoeba fragilis* in patients with suspected enteroparasitic disease in a metropolitan area in Denmark. Clin Microbiol Infect. 2007 Aug;13(8):839-42. doi: 10.1111/j.1469-0691.2007.01760.x
8. **Stensvold CR,** Clark CG, Röser D. Limited intra-genetic diversity in *Dientamoeba fragilis* housekeeping genes. Infect Genet Evol. 2013 Aug;18:284-6. doi: 10.1016/j.meegid.2013.05.003
9. Jokelainen P, Hebbelstrup Jensen B, Andreassen BU, Petersen AM, Röser D, Krogfelt KA, Nielsen HV, **Stensvold CR.** *Dientamoeba fragilis*, a commensal in children in Danish day care centers. J Clin Microbiol. 2017 Jun;55(6):1707-1713. doi: 10.1128/JCM.00037-17
10. **Stensvold CR,** Lebbad M, Victory EL, Verweij JJ, Tannich E, Alfellani M, Legarraga P, Clark CG. Increased sampling reveals novel lineages of *Entamoeba*: consequences of genetic diversity and host specificity for taxonomy and molecular detection. Protist. 2011 Jul;162(3):525-41. doi: 10.1016/j.protis.2010.11.002
11. **Stensvold CR,** Ascuña-Durand K, Chihi A, Belkessa S, Kurt Ö, El-Badry A, van der Giezen M, Clark CG. Further insight into the genetic diversity of *Entamoeba coli* and *Entamoeba hartmanni*. J Eukaryot Microbiol. 2022 Sep 28:e12949. doi: 10.1111/jeu.12949

12. **Stensvold CR,** Nielsen M, Baraka V, Lood R, Fuursted K, Nielsen HV. *Entamoeba gingivalis*: epidemiology, genetic diversity and association with oral microbiota signatures in North Eastern Tanzania. J Oral Microbiol. 2021 May 19;13(1):1924598. doi: 10.1080/20002297.2021.1924598
13. **Stensvold CR,** Winiecka-Krusnell J, Lier T, Lebbad M. Evaluation of a PCR method for detection of *Entamoeba polecki*, with an overview of its molecular epidemiology. J Clin Microbiol. 2018 May;56(5):e00154-18. doi: 10.1128/JCM.00154-18
14. **Stensvold CR,** Lebbad M, Clark CG. Genetic characterisation of uninucleated cyst-producing *Entamoeba* spp. from ruminants. Int J Parasitol. 2010 Jun;40(7):775-8. doi: 10.1016/j.ijpara.2010.03.003
15. Hocke EF, Jamy M, Burki F, Clark CG, **Stensvold CR.** Unravelling the phylogeny of a common intestinal protist: intrageneric diversity of *Endolimax*. Protist. 2022 Oct;173(5):125908. doi: 10.1016/j.protis.2022.125908
16. **Stensvold CR,** Lebbad M, Clark CG. Last of the human protists: the phylogeny and genetic diversity of *Iodamoeba*. Mol Biol Evol. 2012 Jan;29(1):39-42. doi: 10.1093/molbev/msr238
17. **Stensvold CR,** Nielsen SD, Badsberg JH, Engberg J, Friis-Møller N, Nielsen SS, Nielsen HV, Friis-Møller A. The prevalence and clinical significance of intestinal parasites in HIV-infected patients in Denmark. Scand J Infect Dis. 2011 Feb;43(2):129-35. doi: 10.3109/00365548.2010.524659
18. **Stensvold CR,** Lebbad M, Hansen A, Beser J, Belkessa S, O'Brien Andersen L, Clark CG. Differentiation of *Blastocystis* and parasitic archamoebids encountered in untreated wastewater samples by amplicon-based next-generation sequencing. Parasite Epidemiol Control. 2020 May;9:e00131. doi: 10.1016/j.parepi.2019.e00131
19. **Stensvold CR,** Jirků-Pomajbíková K, Tams KW, Jokelainen P, Berg RPKD, Marving E, Petersen RF, Andersen LO, Angen Ø, Nielsen HV. Parasitic intestinal protists of zoonotic relevance detected in pigs by metabarcoding and real-time PCR. Microorganisms. 2021 May 31;9(6):1189. doi: 10.3390/microorganisms9061189

LIST OF TABLES

Table 1. Single-celled eukaryotic genera observed in humans typically referred to as 'parasites' and that can be found in the human gastrointestinal (GI) tract. Genera highlighted in bold font represent genera for which one or more species are commonly acknowledged as pathogenic. 110

Table 2. Current guidelines regarding the use of taxonomic terms for studies of genetic diversity of *Entamoeba*. Reproduced from Jacob et al., 2016 [28]. 111

Table 3. List of consecutive trials of antimicrobial treatment in a Danish patient with irritable bowel syndrome (IBS) with documented *Blastocystis* ST9 and *Dientamoeba fragilis* colonisation (reproduced after [84]). 112

Table 4. Examples of differences in positivity rates reported across the globe in studies using DNA-based methods to detect *Dientamoeba fragilis.* 113

Table 5. Overview of *Dientamoeba fragilis*-specific SSU rDNA sequences (18S only; not ITS sequences with flanking regions generated by Windsor and colleagues) currently available in the NCBI nucleotide database (January 2023) listed according to date of deposition. 114

Table 6. Data from studies involving DNA-based screening of faecal samples from non-human hosts for *Dientamoeba fragilis.* 116

Table 7. Overview of species and ribosomal lineages (RL) of *Entamoeba* based on DNA sequence data available in the NCBI nucleotide database (update of the table included in book chapter by Clark and Stensvold [181] based mainly on references [28, 44, 181, 341]). For more detailed information on ribosomal lineages and conditional lineages, see tables 2 and 8. In this table, 'Ungulates' include elephants for practical reasons. 117

Table 8. *Entamoeba* ribosomal lineages (RL) and conditional lineages (CL) reported until January 2023. 119

Table 9. *Entamoeba coli*-specific SSU rDNA sequences currently available in the NCBI nucleotide database (January 2023) listed according to date of deposition. Subtype information has been provided where possible. A few sequences in the database that cluster with *E. coli* are listed as '*Entamoeba* sp.' and not included in this overview; these include, but may not be limited to, OP796187-OP796190, FN396614, MG256518, MG256520, MG256528, MW133765, MW133768, KY658155, KY658156, KY658157, KY658172, KY658178 and KY658179). Moreover FN396613 and FN396614 are listed in GenBank as *Entamoeba muris*; however, these are likely *Entamoeba coli* ST2. 121

Table 10. *Entamoeba dispar*-specific SSU rDNA sequences currently available in the NCBI nucleotide database (January 2023) listed according to date of deposition. 129

Table 11. *Entamoeba gingivalis*-specific rDNA sequences currently available in the NCBI nucleotide database (January 2023) listed according to date of deposition.* 134

Table 12. *Entamoeba hartmanni*-specific SSU rDNA sequences currently available in the NCBI nucleotide database (January 2023) listed according to date of deposition. A few sequences in the database that are in fact *E. hartmanni* (Figure 10) are listed as '*Entamoeba* sp' (*e.g.,* KF515250, KF515240 and MG25651) or 'Uncultured *Entamoeba* clone' (*e.g.,* JX131943) and not included in the table. Information on subtype was identified where

possible by aligning each sequence with the following near-complete sequences: AF149907 (ST1), KX618191 (ST1), FR686379 (ST1), FR686374 (ST1), FR686377 (ST2), FR686378 (ST3) and FR686376 (ST3). Quite a few of the sequences reflected the 5'-end of the SSU rRNA gene, and the subtype status of most of these remained unresolved; for a few of these, preliminary attempts were made to classify them. 136

Table 13. *Entamoeba moshkovskii*-specific SSU rDNA sequences currently available in the NCBI nucleotide database (January 2023) listed according to date of deposition..... 140

Table 14. Host specificity of *Entamoeba polecki* subtypes 1 to 4 (updated version of the Supplementary Table included in the study by Stensvold et al., 2018 [53])..... 149

Table 15. *Entamoeba polecki*-specific SSU rDNA sequences currently available in the NCBI nucleotide database (January 2023) listed according to date of deposition. In the second column, the entries for 'Organisms' were those originally entered in the NCBI database. *Entamoeba polecki* includes *E. polecki* subtypes 1 to 4, whereas *Entamoeba chattoni* and *Entamoeba struthionis* are redundant names for *E. polecki* ST2 and *E. polecki* ST3, respectively. 150

Table 16. The host species of *Entamoeba bovis* identified so far based on DNA sequence data from the NCBI nucleotide database (as of January2023). 158

Table 17. *Endolimax*-specific SSU rDNA sequences currently available in the NCBI nucleotide database (January 2023) listed according to date of deposition. 159

Table 18. *Iodamoeba*-specific SSU rDNA sequences currently available in the NCBI nucleotide database (January 2023) listed according to date of deposition. 161

LIST OF FIGURES

Figure 1. Schematic overview of the types of intestinal parasites found in humans. Helminths are multi-cellular, and adult individuals are macroscopically discernible. Meanwhile, the remaining organisms are single-celled and not macroscopically discernible. The group 'Other' includes genera such as *Blastocystis*, *Enterocytozoon*, and *Encephalitozoon*. 19

Figure 2 (previous page). Phylogenetic relationship between the subtypes of *Blastocystis* acknowledged to date. For higher resolution, please visit file repository on GitHub (https://github.com/Entamoeba/DMSc-Thesis/blob/main/Blastocystis%20Tree%20GitHub.pdf). The evolutionary history was inferred using the Neighbor-Joining method (Kimura 2-parameter method). The analysis involved 96 nt sequences. There were a total of 1,189 positions in the final dataset. The scale bar indicates nt substitutions per site. 37

Figure 3. Analysis of intra-subtype diversity of *Blastocystis* ST4 SSU rDNAs. Subtype 3 was used as outgroup. There were a total of 1,734 positions in the final dataset. Two clades are appreciated, one of which (Clade 1) is genetically homogenous and holds sequences from several host groups, and the other of which (Clade 2) is genetically heterogeneous and holds sequences almost exclusively from rodents. The scale bar indicates nt substitutions per site. 40

Figure 4, A and B. Example of how *Entamoeba* phylogeny has developed over the past couple of decades. The analysis comprises complete or near-complete SSU rDNA *Entamoeba* sequences. Example A is a reproduction of the tree included in the work by Clark and colleagues from 2006 [185]. In example B, all ribosomal lineages and subtypes reported to date (January 2023) were included (reference sequence depository available here: http://entamoeba.lshtm.ac.uk/ref.entamoeba.txt). The sequences highlighted in bold font are sequences that were published for the first time in the articles shortlisted for this thesis. The tree in part B was generated for the present thesis. It used distance-based analysis (Neighbour-Joining algorithm with 1,000 bootstraps) of 1,526 positions in an alignment of 37 nt sequences. The scale bar indicates nt substitutions per site. 51

Figure 5 (previous page). Intraspecies variability of *Entamoeba coli* based on 104 DNA sequences representing the 5'-end of the SSU rRNA gene retrieved from the NCBI database. The *E. coli*-specific sequences obtained from non-human hosts are indicated in boldface type. Non-human *E. coli* hosts are indicated in parentheses after the NCBI database ID accession number. Three subtypes are acknowledged (ST1–ST3). The Neighbor-Joining method was used. Evolutionary distances were computed using the Kimura 2-parameter method. All ambiguous positions were removed for each sequence pair (pairwise deletion option). There were 244 positions in the final dataset. Only bootstrap values >70 are shown. The scale bar indicates nt substitutions per site. The figure is reproduced from the work by Stensvold et al., 2022 [44]. 54

Figure 6. Neighbor-joining analysis (1,000 bootstraps) of the sequence MW026738 (highlighted in boldface) with reference sequences reflecting all three suggested subtypes of *Entamoeba coli* and AF149913 (*Entamoeba polecki*) as outlier. The analysis could suggest that MW026738 represents a novel subtype, which in that case could be referred to as ST4. There were a total of 554 positions in the final dataset. The scale bar indicates nt substitutions per site. 55

Figure 7. Phylogenetic analysis of SSU rDNA sequences of *Entamoeba histolytica*, *Entamoeba dispar*, and *Entamoeba nuttalli.* The tree was generated by Neighbour-Joining method-based analysis of 1,658 unambiguously aligned nt positions, corresponding to about ~90% of the SSU rRNA gene. Information on host and geography is provided where available in GenBank. Three clades can be appreciated: The upper clade is separated from the remaining sequences by a bootstrap support of 86 and holds *E. nuttalli* sequences (although some were deposited in

GenBank as '*E. histolytica*'). The second clade holds *E. histolytica* (*sensu stricto*) sequences and is supported by a bootstrap value of 100. The third clade is also supported by a bootstrap value of 100 and holds sequences specific to *E. dispar*. The *E. dispar* clade again holds two clusters, which could indicate the presence of two subtypes of *E. dispar.* The genetic distance between the two clusters is up to about 0.5%. The scale bar indicates nt substitutions per site. 59

Figure 8. Phylogenetic analysis of rDNA sequences of *Entamoeba gingivalis* reflecting the 3'-end of the 18S + the ITS1 region. Three subtypes can be appreciated. The tree was generated by neighbour-joining analysis of 348 unambiguously aligned nt positions. The evolutionary history was inferred using the Neighbor-Joining method and the evolutionary distances were computed using the Kimura 2-parameter method. The scale bar indicates nt substitutions per site. 61

Figure 9. Phylogenetic analysis of SSU rDNA sequences of *Entamoeba gingivalis.* The tree was generated by neighbour-joining analysis of 330 unambiguously aligned nt positions, corresponding to about ~20% of the SSU rRNA gene. Sequences without GenBank accession numbers are sequences obtained in our study from Tanzania [43].Two clades can be appreciated; the top clade with the majority of the sequences represents ST1, and the lower clade (with three sequences, which are all reference sequences) represents ST2. The scale bar indicates nt substitutions per site. 62

Figure 10. Phylogenetic analysis of *Entamoeba hartmanni*-specific sequences generated in the study by Stensvold *et al.* [44] revealing the existence of three subtypes (ST1–ST3). The Neighbor-Joining method was used. Evolutionary distances were computed using the Kimura 2-parameter method. The analysis involved 43 nt sequences. There was a total of 411 positions in the final dataset. The region covered corresponded to the middle part of the SSU rRNA gene (as opposed to the sequences included in Figure 11, which reflected the 5′-end of the gene, and which does not enable differentiation of subtypes). Only bootstrap values >70 are shown. Sequences generated in the present study are highlighted in boldface and were all from humans. The scale bar indicates nt substitutions per site. Abbreviations used in parentheses for sample origin are as follows: A, Africa; E, Europe; SA, South America. 65

Figure 11. A number of ~500 bp-long SSU rDNA sequences reflecting Entamoebas found in mountain gorillas in Uganda were deposited in GenBank by Nolan and colleagues in 2017 [197]. These cluster with *E. hartmanni*, but since the sequences reflect the 5'-end of the SSU rRNA gene, it was not possible to identify which *E. hartmanni* subtype they belong to [44]. The same applies to the sequences JX131936 and JX131943 found in a study by Hamad and colleagues [243]. The scale bar indicates nt substitutions per site. The tree was reproduced from the study by Stensvold et al., 2022 [44]..... 66

Figure 12. Genetic diversity of *Entamoeba moshkovskii* based on sequences available in the NCBI nt database covering ~80% of the SSU rRNA gene or more (n = 26). Two major clades can be appreciated. There were a total of 1,362 positions in the final dataset. The evolutionary history was inferred using the Neighbor-Joining method (Kimura 2-parameter method). The scale bar indicates nt substitutions per site. 68

Figure 13. Phylogenetic analysis of *Entameoba polecki* sequences recently published in GenBank (boldface type) together with a selection of *E. polecki* reference sequences representing the four subtypes acknowledged to date (1,000 bootstraps). The Neighbor-Joining method was used. MK801450 is likely a chimaera (see text for details), which could explain why it appears as a deep branch of ST3. In Figure 14, the analysis was repeated without MK801450 to try to obtain better resolution of OP9019601 (sequence from dog) and OP753638 (sequence from cat).There were a total of 562 positions in the final dataset. The scale bar indicates nt substitutions per site. Only bootstrap values >60 are shown. Information on host and geographical origin is provided where known. 70

Figure 14. Neighbor-joining analysis of two *Entameoba polecki* sequences recently published in GenBank (boldface type) together with a selection of *E. polecki* reference sequences representing the four subtypes acknowledged to date (1,000 bootstraps). The phylogenetic analysis is similar to that in Figure 13, apart from the fact that MK801450 was not included, since it was likely a chimaera (see text for details). The present analysis was repeated without MK801450 to try to obtain better resolution of OP9019601 (sequence from dog) and OP753638 (sequence from cat). There were a total of 562 positions in the final dataset. The scale bar indicates nt substitutions per site. Only bootstrap values >70 are shown. Information on host and geographical origin is provided where known. 71

Figure 15. Maximum likelihood phylogeny of *Endolimax* and relatives (including *Iodamoeba* RL 1 and RL2), reconstructed from an SSU rDNA alignment consisting of 21 taxa and 2,067 positions. Maximum likelihood bootstrap values and Bayesian posterior probabilities are shown in that order on each bipartition. GenBank accession numbers are indicated in parentheses. The scale bar indicates nt substitutions per site. The sequences that were generated in our study [32] are indicated with a star; sequences from sewage have the prefix SW. Abbreviation: RL, ribosomal lineage. (Adapted from [32]). 75

SUMMARY IN DANISH

Klinisk mikrobiologi involverer påvisning og differentiering af primært bakterier, virus, parasitter og svamp hos patienter med infektioner.

Op til flere milliarder mennesker kan være bærere af en eller flere arter af encellede tarmparasitter som ofte påvises i klinisk mikrobiologiske laboratorier; alligevel er vor viden om disse organismers indvirkning på den globale folkesundhed stærkt begrænset. Slægterne *Blastocystis*, *Dientamoeba*, *Entamoeba*, *Endolimax* og *Iodamoeba* omfatter hovedparten af disse tarmparasitter, hvis livscyklus, i modsætning til andre encellede tarmparasitter (f.eks. mikrosporidier og sporozoer), formentlig ikke involverer tarm-invasive stadier der kan resultere i patologiske processer og dermed sygdom (undtagelsen er *Entamoeba histolytica*). Alle fem slægter er parasitter i den forstand, at de benytter en vært for at fuldføre livscyklus; dog anses de af mange som værende af begrænset klinisk relevans og burde måske snarere kaldes ”endobionter”, eller ”endosymbionter”, dersom de måtte have sundhedsbeskyttende funktioner. De artikler der er inkluderet i denne afhandling eksemplificerer noget af det arbejde der har resulteret i, at det nu synes mere relevant at studere omtalte organismer i en folkesundhedsmæssig og tarmøkologisk sammenhæng end i en klinisk mikrobiologisk sammenhæng.

Undersøgelser af indvirkningen af tarmparasitter på sundhed og sygdom kræver præcise diagnostiske værktøjer, herunder især DNA-baseret teknologi som PCR og sekventering, samt akkurate referencedatabaser. ’Small-subunit’ (SSU) ribosomale RNA (rRNA) gener, som findes i alle pro- og eukaryote organismer, bruges i dag flittigt som taksonomiske markører. DNA-baserede metoder kan udvikles til genetisk karakterisering af mikroorganismer og dermed give information om species/subtype/genotype/osv. Metagenom og metabarcoding analyser (anvendelsen af lav-specifik PCR koblet til ‘next-generation’ sekventering) kan give information om co-infektion/co-kolonisation med diverse species og muliggør screening for genetisk diversitet i selv komplekse matricer samt afdækning af kryptisk genetisk diversitet (se nedenfor).

Ved at udvikle og implementere følsomme og præcise DNA-baserede diagnostiske værktøjer og typningsmetoder primært baseret på SSU rRNA genet har vi øget indsigten i diversiteten, fordelingen og den folkesundhedsmæssige betydning af encellede luminale tarmparasitter (ELTP). Med disse redskaber har vi vist, at slægterne *Blastocystis* og *Dientamoeba* er langt hyppigere forekommende end antaget for blot 10–15 år siden, hvor hypoteser om forekomst beroede på de data der var til rådighed fra den traditionelle parasitologiske diagnostik, typisk mikroskopi, eller hvor der slet ingen data var. Således har vi vist, at de fleste børn i daginstitutioner i Danmark bliver smittet med *Dientamoeba*. Ligeledes har vi vist, at større børn i Nigeria stort set alle har *Blastocystis*.

Encellede luminale tarmparasitter kan forekomme hos patienter med diare og funktionelle eller inflammatoriske mave-/tarm-lidelser. Dog tegner der sig et klart billede af, at ELTP forekommer langt hyppigere hos tarm-raske individer end hos patienter med mave-tarm-symptomer, og koblet til den forskning vi har udført vedrørende sammenhæng mellem ELTP og tarmbakterier tyder det på, at kolonisation med disse parasitter primært ses hos individer med en "sund tarmflora" (eubiose). Disse fund bør i fremtiden informere og inspirere forskningsindsatser med fokus på udnyttelse af ELTP som biomarkører, og det bør undersøges, i hvilket omfang manipulation med ELTP kan afstedkomme ændringer i "tarmflora" og dermed finde anvendelse som f.eks. probiotika.

Dersom der kan være tale om infektion med ELTP, mangler vi stadig værktøjer til at skelne kolonisation fra infektion. Det har længe stået klart, at tilsyneladende morfologisk identiske parasitter kan have forskellig klinisk betydning og være genetisk forskellige, hvilket kaldes for "kryptisk genetisk diversitet". Et eksempel er *E. histolytica* som på cyste-form ikke lader sig adskille fra *Entamoeba dispar*. Men hvor *E. histolytica* kan være invasiv og give anledning til f.eks. amøbedysenteri og amøbiasis, anses *E. dispar* af de fleste for uskadelig. Denne indsigt ledte os til at undersøge den genetiske diversitet yderligere hos *Entamoeba* og andre slægter af ELTP. Hvis lignende – eller endda mere udtalt – diversitet kunne påvises hos *Blastocystis*, *Dientamoeba*, *Endolimax* og *Iodamoeba*, kunne disse forskelle muligvis forklare forskelle i symptom-status og kolonisation versus infektion. Til trods for endda markant intragenetisk diversitet hos nogle af disse (se nedenfor), har vi til dato fundet begrænset holdepunkt for en sådan teori, men der er brug for flere og mere omfattende studier. For *Dientamoeba* er der tale om en næsten klonal ekspansion af den ene af de to genotyper der kendes i dag; denne genotype synes endda at være altdominerende. Omvendt ses der udtalt genetisk diversitet hos *Blastocystis*, og der er til dato fundet mindst 30 arter, de såkaldte 'subtyper'. Vi, og mange andre, har søgt at identificere om visse subtyper er særligt associeret til udvikling af tarm-symptomer, men der er ikke stærk evidens herfor. Dog ved vi nu, at subtyperne 1-4 udgør omkring 95% af *Blastocystis* kolonisation hos mennesker, og vi har vist, at mennesker der smittes med zoonotiske subtyper (fx ST6, ST7 og ST8) typisk har symptomer.

Forbløffende genetisk variation har vi også påvist hos andre ELTP. Dette har bl.a. ført til anerkendelsen af *Iodamoeba bütschlii*, *Endolimax nana,* og flere *Entamoeba* arter som specieskomplekser, hvor hver art bør betragtes som et kompleks af arter (kaldet subtyper eller ribosomale linjer) med overlappende morfologi. Hvor *E. histolytica* og *E. dispar* adskilles ved kun 1%–2% diversitet på tværs af SSU rRNA genet, så er der f.eks. godt 10% og 30% forskel på ribosomale linjer af *E. coli* henholdsvis *I. bütschlii,* hvilket udfordrer de species-begreber vi har opereret med indtil nu. Min forskning har ført til anerkendelsen af tre ribosomale linjer af både *E. coli* og *E. hartmanni*, samt to ribosomale linjer af *E. nana* og *I. bütschlii*. Desuden har vi fundet mindst én ny ribosomal linje af *E. moshkovskii*.

Genetisk karakterisering af tarmparasitter indsamlet fra forskellige værtsarter (mennesker, ikke-humane primater, andre pattedyr, fugle, *etc.*) kan øge indsigten i omfanget af smitte mellem dyr og mennesker. Vi har vist, at grise kan være vært for arter/ribosomale linjer af ELTP der er almindelige hos mennesker, f.eks. *Entamoeba hartmanni* og *I. bütschlii.* Mange andre arter af større pattedyr er almindelige værter for f.eks. *Blastocystis* og *Entamoeba*; dog kan de arter/genetiske varianter der findes af *Blastocystis* og *Entamoeba* hos disse dyr være forskellige fra dem vi ser hos mennesker. Dette tyder på, at disse slægter har adapteret sig til deres respektive værter over en forholdsvis lang periode, med en relativ høj værtspecificitet til følge. For *Blastocystis* har vi endda vist, at selvom en given subtype (ST) kan forekomme hos mere end en værtsart, så er det muligt at påvise kryptisk værtsspecificitet på stamme-niveau. Dette vil sige, at selvom mennesker og ikke-humane primater begge kan være koloniseret med f.eks. *Blastocystis* ST3, så er det i høj grad værts-specifikke stammer af ST3 der cirkulerer i de to værtspopulationer. Det er også muligt, at hvor *E. coli* ST1 er tilpasset mennesker, så er *E. coli* ST2 muligvis en art der har ikke-humane primater og gnavere som reservoir.

Der står efterhånden klart, at ELTP er almindelige fund hos baggrundsbefolkningen, selv i Danmark. Vi har endnu ikke har kunnet skelne kolonisation fra evt. infektion, og der skulle være belæg for at mene, at individer med ELTP sjældent har brug for eller nytte af klinisk intervention for disse, herunder medikamentel behandling. Måske er det endda sådan, at man bør forsøge at undgå at fjerne disse organismer fra tarmen, hvis de først har etableret sig. Dog bør der foretages yderligere studier der belyser betydningen af den udtalte genetiske diversitet hos visse ELTP i forhold til smitte og klinisk betydning. Fremtidens forskning i ELTP bør også inkludere studier der kan belyse de faktorer der favoriserer kolonisation med ELTP og hvilken betydning ELTP har for værtens tarmøkologi, metabolisme og generelle sundhedstilstand. Endelig, da mennesker og dyr deler nogle af disse slægter/arter, og da nogle protozoer har betydning for tarmfunktionen hos f.eks. drøvtyggere, kunne det være interessant at studere disse parasitter i såvel domesticerede som vilde dyr yderligere for at lære mere om parasitternes indvirkning på sundhed og sygdom, herunder vurdere, hvorvidt nogle ELTP er endosymbionter.

SUMMARY IN ENGLISH

Clinical microbiology involves the detection and differentiation of primarily bacteria, viruses, parasites and fungi in patients with infections.

Billions of people may be colonised by one or more species of common luminal intestinal parasitic protists (CLIPPs) that are often detected in clinical microbiology laboratories; still, our knowledge on these organisms' impact on global health is very limited. The genera *Blastocystis*, *Dientamoeba*, *Entamoeba*, *Endolimax*, and *Iodamoeba* comprise CLIPPs species, the life cycles of which, as opposed to single-celled pathogenic intestinal parasites (*e.g.,* microsporidia and sporozoa), do probably not include gut-invasive stages that could result in pathological processes and thereby disease (except for *Entamoeba histolytica*). All five genera are parasites in the sense that they use a host to complete their life cycle; still, by many specialists, these are considered to be of limited clinical relevance and could possibly be referred to as 'eukaryotic endobionts' or even 'endosymbionts', in case they would have health-protective effects. The articles included in this thesis exemplify the work and the data that support the view that it might be more relevant to study these genera in a public health and gut ecology context than in a clinical microbiology context.

Essential to investigating the impact of intestinal parasites on health and disease are accurate diagnostic tools, including DNA-based technology such as PCR and sequencing, plus accurate reference databases. Small subunit (SSU) ribosomal RNA (rRNA) genes consistently present in both pro- and eukaryotic organisms are today avidly used as taxonomic markers. DNA-based methods have been developed for genetic characterisation of microorganisms and provided data on species/subtypes/genotypes/*etc.* Metagenomics and metabarcoding (the use of low-specific PCR coupled with next-generation sequencing) can provide information on co-infection/co-colonisation with other organisms and enable screening for genetic diversity, even in complex matrices.

By developing and implementing sensitive and specific DNA-based diagnostic tools and typing assays primarily based on the SSU rRNA gene, we have increased insight into the diversity, distribution and significance of CLIPPs. With these tools, we have shown that the genera *Blastocystis* and *Dientamoeba* are far more common than previously thought. Only 10–15 years ago, hypotheses on their distribution typically relied on data generated by traditional parasitological diagnostic methods, such as light microscopy. Hence, we have shown that most older children in Nigeria host *Blastocystis*, and that most children in day-care institutions in Denmark, if not all, get colonised by *Dientamoeba* at some point.

Single-celled non-pathogenic intestinal parasites can be hosted by patients with diarrhoea and functional or inflammatory bowel diseases. However, emerging data appear to suggest that CLIPPs are generally more

common in gut-healthy individuals than in patients with gastrointestinal symptoms. The research we have carried out on associations between CLIPPs and gut bacteria suggests that colonisation with these parasites is seen primarily in individuals with a healthy 'gut flora' (eubiosis). This observation should prompt future research projects focusing on the use of CLIPPs as biomarkers, and it should be investigated to which extent manipulation with CLIPPs could lead to changes in the gut flora and thereby be used as probiotics.

In the event that it makes sense to speak of 'infection' by CLIPPs, we still lack tools to differentiate between colonisation and infection. We have known for decades that morphologically similar parasites can differ in terms of clinical impact and be genetically distinct, a feature that we refer to as 'cryptic genetic diversity'. One example is *Entamoeba histolytica*, which cannot be differentiated from *Entamoeba dispar* by cyst morphological features. However, whereas *E. histolytica* can be invasive and give rise to amoebic dysentery and amoebiasis, *E. dispar* is by most specialists considered non-invasive and generally non-pathogenic. This insight led us to investigate genetic diversity among other species of *Entamoeba* as well as other CLIPPs genera. If we could demonstrate similar – or higher – degrees of diversity within *Blastocystis*, *Dientamoeba*, *Endolimax*, and *Iodamoeba*, these differences might be key to explaining differences in parasite phenotype and thereby differences in the ability of the parasites to cause symptoms. Despite the disclosure of striking genetic diversity among some CLIPPs, we have found little support for such theories; however, more studies are needed. As for *Dientamoeba*, we have observed a more or less clonal expansion of one of the two genotypes known to exist, and this genotype appears to have global predominance. In contrast, extensive genetic diversity is observed between and within subtypes of *Blastocystis*: to date, more than 30 species, the so-called subtypes, have been acknowledged. We, and many others, have sought to identify whether one or more of these subtypes could be linked to the development of intestinal symptoms, but there is little evidence to support this hypothesis. We know that subtypes 1–4 reflect about 95% of *Blastocystis* colonisation in humans, and we have shown that individuals with zoonotic subtypes (*e.g.,* ST6, ST7, and ST8) might typically experience symptoms.

We have disclosed astonishing genetic variation among other CLIPPs, which has led to the recognition of *Iodamoeba bütschlii*, *Endolimax nana, Entamoeba coli* and *Entamoeba hartmanni* as species complexes, where each species should be regarded as a complex of species (referred to as 'subtypes' or 'ribosomal lineages') with overlapping morphology. And where *E. histolytica* and *E. dispar* differ by only 1%–2% diversity across the SSU rRNA gene, we have observed up to at least 10% and 30% genetic difference among ribosomal lineages within *E. coli* and *I. bütschlii*, respectively, challenging species concepts currently applied. Our research has resulted in the recognition of three ribosomal lineages within both *E. coli* and *E. hartmanni*, as well as two ribosomal lineages of *E. nana* and *I. bütschlii*. Moreover, we have discovered a new lineage of *Entamoeba moshkovskii*.

Molecular characterisation of intestinal parasites collected from different host species (humans, non-human primates, other mammals, birds, *etc.*) can help identify opportunity for transmission between human and non-human hosts. We have shown that pigs can host a few species/lineages that can readily colonise humans, such as *Entamoeba hartmanni* and *I. bütschlii*. Many other species of larger mammals are common hosts of *Blastocystis* and *Entamoeba*. However, for the two latter genera, the species/genetic variants observed in non-human hosts are typically different from those observed in humans, which could indicate that many species of CLIPPs have adapted to their respective hosts over a long period, resulting in relatively high host specificity. For *Blastocystis*, we have shown that even though a given subtype may be found in more than one host species, it is possible to demonstrate cryptic host specificity at allele level. For instance, even though both human and non-human primates can be colonised by ST3, host species-specific strains of ST3 circulate within these two host populations. With regards to *E. coli*, it is possible that ST1 has adapted to human hosts, while *E. coli* ST2 has adapted to a broader host range, including non-human primates and rodents.

It has become clear that CLIPPs are common colonisers of the human background population, and even though we cannot disprove the existence of infection by any of these, it should be reasonable to consider clinical and medical intervention redundant in most cases. Perhaps it might even be so that one should try not to eradicate these organisms from the gut when first established. However, more studies are warranted to elucidate the significance of the pronounced genetic diversity observed in some CLIPPs with regards to transmission patterns and clinical significance. Future research in CLIPPs should also include studies that can elucidate those factors that favour colonisation with CLIPPs and what role CLIPPs have in host gut ecology, metabolism and overall health condition. Finally, since human and non-human hosts share these parasitic genera, and since some protozoa possibly contribute to overall gut function in ruminants, it would be interesting to study these in domesticated and wild animals to learn more about the role of these parasites in health and disease, including investigations into whether some CLIPPs might be endosymbionts.

1. INTRODUCTION

1.1. Common Luminal Intestinal Parasitic Protists (CLIPPs) – Overall Classification

Any non-fungal eukaryotic organism living in/on and completing at least part of its life cycle in humans could be referred to as a 'parasite'. Parasites comprise those affecting the skin (ectoparasites) and those living inside the body (endoparasites). Endoparasites can for practical reasons be divided into tissue, blood and digestive-tract parasites. Parasites may be found in the human digestive tract from the oral cavity to the distal colon and anal canal. A simplistic but potentially useful categorization of organisms parasitising the human intestine is provided in **Figure 1**. Overall, intestinal parasites of humans are either multicellular (metazoa; the 'worms') or single-celled (parasitic protists, most of which are protozoa). Previously considered protozoa, the genera *Enterocytozoon* and *Encephalitozoon* are now taxonomically classified as fungi; in the clinical microbiology laboratory (CML), however, these are still mostly dealt with in the parasitology section of the CML. The life cycle of intestinal microsporidia largely resembles that of sporozoa, such as *Cryptosporidium*, with similar transmission modes and involving an intracellular life cycle stage and the shedding of spores in faeces that can be detected by for instance staining of faecal concentrates or DNA-based methods. Also *Blastocystis* is an "outsider", as it belongs to the group of Stramenopiles (heterokonts). An overall breakdown of the single-celled intestinal parasitic genera commonly or less commonly (depending on the population examined) observed in human stool samples is provided in **Table 1.** Importantly, many of these genera can be encountered in non-human hosts.

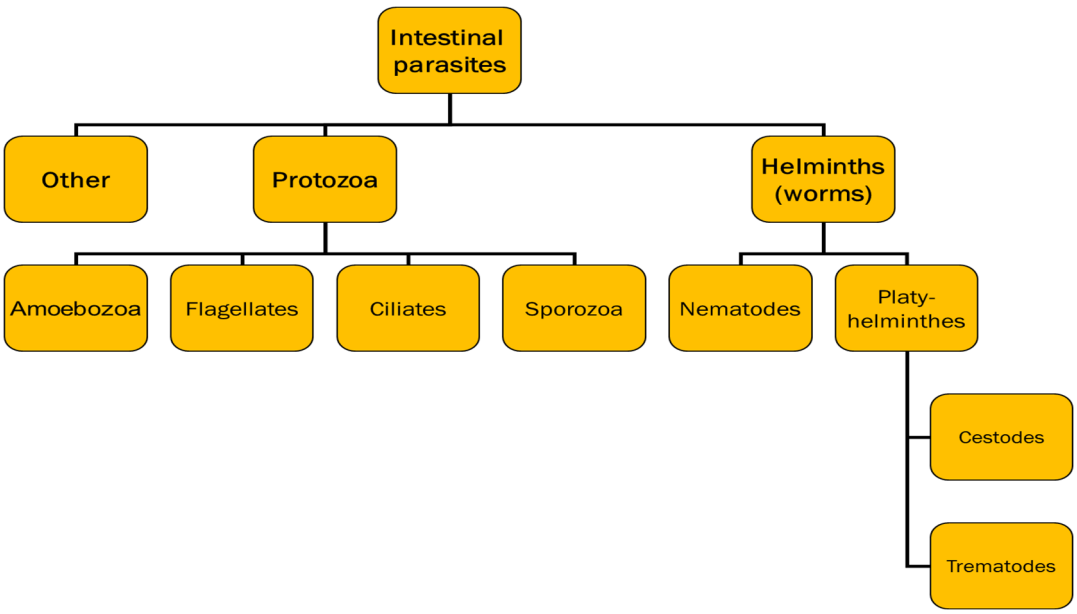


Figure 1. Schematic overview of the types of intestinal parasites found in humans. Helminths are multi-cellular, and adult individuals are macroscopically discernible. Meanwhile, the remaining organisms are single-celled and not macroscopically discernible. The group 'Other' includes genera such as *Blastocystis*, *Enterocytozoon*, and *Encephalitozoon*.

While some single-celled intestinal parasites are mostly observed in patients with gastrointestinal symptoms, the reverse is true for others. For instance, with respect to individuals tested in Denmark, genera such as *Cryptosporidium*, *Giardia* and *Enterocytozoon* are only very rarely found in asymptomatic individuals, whereas these are increasingly being acknowledged as causes of (outbreaks of) diarrhoea and gastroenteritis in Scandinavia [1-8]. Meanwhile, several other single-celled intestinal parasite genera appear to be much more common in gut-healthy individuals than in patients with intestinal symptoms and individuals with underlying diseases; these parasites are the ones that are the focus of this thesis.

In the scientific literature, the term 'parasite infection' may often be encountered. While it makes sense to define cryptosporidiosis as a disease caused by a *Cryptosporidium* infection due to the invasive stage of the *Cryptosporidium* life cycle, the pathology caused including the triggering of an immune response, it may be misleading to refer to, for example, a '*Blastocystis* infection', in the event that no invasive stage, pathology, or even immune reaction is involved during intestinal colonisation. Hence, it may be practical to differentiate between 'infection' and 'colonisation', where the latter does not typically activate the immune system or involve any pathology. Hence, it appears relevant to speak of 'cryptosporidiosis', whereas the term 'blastocystosis' would typically make little sense. This line of thinking also underlies the reason why I coined the term 'common luminal intestinal parasitic protists' (CLIPPs), where the word 'luminal' is meant to indicate the absence of gut-invasive properties, and where 'protists' were chosen rather than 'protozoa', which term would exclude *Blastocystis* [9].

The life cycles of most CLIPPs involve both a trophozoite(-like) stage and a cyst stage; however, for some (*e.g., Entamoeba gingivalis* and *Dientamoeba fragilis*), cyst stages remain to be identified and confirmed, respectively. Transmission occurs via the faecal-oral route.

Although perceived as non-pathogenic by most specialists, it has been customary in the CML to report these parasites when found in faecal samples, as the presence of these organisms indicate exposure to faecal contamination and prompt further investigations for pathogenic organisms.

1.2. Detection and Differentiation of CLIPPs

Except for *Blastocystis*, the organisms of study in this thesis were described in remarkable detail by Clifford Dobell (1886–1949), a British protozoologist. Most of Dobell's observations were published during the 1910s and 1920s and had vast implications for future parasitological evaluation of faecal samples. Since then, classical examination for single-celled parasites in human stool samples has involved concentration of parasitic elements and removal of fat and debris from the samples to enable detection of (oo)cysts of protozoa and ova of helminths, and microscopy of fresh faeces and/or permanent staining of fixed faecal material has been used to detect trophozoites of protozoa and stages of *Blastocystis*. Identification of protozoa such as

Entamoeba involves morphological description of any cysts observed, including information on size, number of nuclei, position of karyosome, presence/absence of peripheral chromatin in the nuclei and chromatid bodies among other features, and/or description of trophozoite stages, which may also exhibit morphological hallmarks. Indeed, data on morphology combined with information on the sampled host have been used to name parasites using the classical binomial nomenclature. However, the use of DNA-based methods has made it clear that detection and differentiation of single-celled intestinal parasites based on light microscopy alone has at least two major limitations: The first limitation has to do with the limited sensitivity of microscopy-based methods compared with DNA-based methods [10, 11] which is an important point when the aim is to separate colonised from non-colonised individuals. The second limitation pertains to the fact that morphologically similar organisms may exhibit extensive genetic diversity, a feature referred to as 'cryptic genetic diversity' (see below).

Over the past two decades, direct diagnosis of intestinal parasites has changed from relying predominantly on microscopy, including staining of faecal concentrates, to relying primarily on DNA-based methods, at least in the modern CML [11, 12]. The incentive for the use of DNA-based methods has been highlighted by Stensvold et al. [10] and Verweij and Stensvold [11] and includes increased sensitivity, automatability, standardisability, and the opportunity easily to store and use positive DNAs for molecular typing where needed. DNA-based methods can be tailored to specific needs and include singleplex, duplex and multiplex real-time PCR assays, and, more lately, metabarcoding assays, just to mention some options that can be applied to genomic DNA extracted either directly from a sample or after any pre-DNA extraction step that would aim to enhance the sensitivity and overall quality of the assay. PCR with Sanger sequencing has been used extensively for differentiation of parasitic species or subspecies/subtypes/genotypes, depending on the parasite in question.

To this end, it should be mentioned that the nuclear small subunit (SSU) ribosomal RNA (rRNA) gene has over time proved a quite robust taxonomic identifier for intestinal parasites. The nucleotide database at the National Center for Biotechnology Information (NCBI) (referred to in this thesis as either 'NCBI nucleotide database', 'NCBI database' or 'GenBank') and other databases are continuously being updated with sequence data from activities involving Sanger sequencing, but also next-generation sequencing (NGS). Metabarcoding assays have been developed and already used extensively. One such example is the metabarcoding assay in place at Statens Serum Institut (SSI), which enables detection and differentiation of nuclear small subunit rRNA genes (also known as 16S and 18S genes) of bacteria, parasites and fungi, which means that this one-fits-most approach can be used to detect a variety of organisms simultaneously in anyone sample, although with varying sensitivity [13-15]. This 'meta-ribosomalomics' approach is particularly useful to identify mixed colonisation by genetically related species or subtypes that cannot be differentiated otherwise (*e.g.,* by microscopy or even other NGS-based assays [16]) and to assist in mapping the 'eukaryome' of selected host species [17].

Given the likelihood of CLIPPs being confined to the gut lumen not triggering immune responses observed for parasites that cause pathology or those that at least are invasive (*e.g., Toxoplasma*), one would not expect serology to be a relevant diagnostic modality for CLIPPs detection. To my knowledge, *Blastocystis* is the only CLIPP where serological diagnosis has been attempted [18]. Hence, indirect detection methods appear to be of little relevance to CLIPPs.

1.3. Cryptic Genetic Diversity

Molecular characterisation of parasites is useful in epidemiological surveillance, including outbreak investigations, and for identification of lineages that differ in prevalence among healthy and diseased individuals, those that differ geospatially, and for revealing transmission patterns and host reservoirs. Within the field of intestinal parasite research, the relevance of molecular characterisation has been highlighted by the observation of cryptic genetic diversity within *Entamoeba*. Organisms that share identical morphological features but that differ genetically (and maybe also phenotypically) to an extent where it could be argued that they reflect different species and not only different genotypes may be referred to as 'cryptic species' [19]. In the early nineties, it was finally confirmed that some morphologically similar *Entamoeba* parasites were genetically different, leading to the separation of *E. dispar* from *E. histolytica* [20, 21]. Since then, *E. dispar* has been considered largely non-pathogenic, while *E. histolytica* is known for its dramatic pathological potential [22-24]. This recognition has had important consequences for the diagnosis and management of patients with *Entamoeba*-positive stool samples.

Entamoeba cysts with eight nuclei and with overlapping size ranges may be found in faeces from both primates and rodents; however, the cysts from rodents may genetically differ from those of primates by more than 17% across the entire SSU rRNA gene, and even within *Entamoeba coli* – which is the octonucleated cyst producer observed in humans – up to at least 12% divergence has been observed (see section 2.3.1). Even higher differences (~30%) have been observed for morphologically similar cysts of *Iodamoeba* (see section 2.5); to put this into perspective, some species – for instance *E. dispar* and *E. histolytica* differ by only 1%-2% across the entire SSU rRNA gene (Novati and colleagues mentioned 1.7% nucleotide (nt) substitutions [25]), and several species of *Cryptosporidium* differ by only 2%-4%.

Another example has to do with the species name "*Blastocystis hominis*", a term now considered invalid [26] and which was applied to *Blastocystis* identified in stool samples. However, *Blastocystis* with indistinguishable morphological features can be found in humans and a vast variety of non-human hosts, and studies of the genetic makeup of these morphologically similar organisms indicates the existence of multiple species (currently referred to as "subtypes"; see section 2.1). Hence, molecular methods enable a higher and more accurate discriminatory ability compared with microscopy, and obviously, molecular methods are

particularly useful for studying organisms for which cyst morphology data are not available (*e.g., D. fragilis* and *E. gingivalis*).

On the other hand, basing taxonomic inferences on molecular data only also has its limitations, which is why it has become common to use alternatives to the traditional Latin binomial nomenclature to delineate species at least for a preliminary period until data from sampling of relevant hosts are sufficient to make robust inferences on host specificity and until there are morphological data that can be paired with the DNA data. Terms commonly in use are 'subtypes' (*e.g., Blastocystis* and some species of *Entamoeba*) and 'ribosomal lineages' (*e.g., Entamoeba, Iodamoeba* and *Endolimax*).

There are no rules set in stone as to how genetically different organisms have to be in order to be considered two different species/lineages. For *Blastocystis*, we now recommend that a DNA sequence can be a candidate for a new subtype number if it covers >80% of the entire SSU rRNA gene and differs by >4% from previously sequenced complete *Blastocystis* SSU genes [27]. For practical reasons, colleagues provided definitions of species, subtypes, ribosomal lineages, and conditional lineages for studies of genetic diversity of *Entamoeba* [28] (**Table 2**).

Importantly, terminology should be practical and pragmatic, and it is beyond the scope of the present research to venture into more theoretical discussions on species concepts and the more theoretical aspects on which taxonomical inferences are based.

1.4. Molecular Characterisation of Parasites

Molecular characterisation of parasite genera from human and non-human hosts can help us identify whether transmission between host species, for instance between non-human and human hosts, might take place. In other words, molecular characterisation can help us identify whether zoonotic transmission is likely for any parasite species in question by comparing parasite data from human and non-human hosts. An intestinal parasite for which molecular characterisation has proven particularly useful in order to delineate patterns and the extent of zoonotic transmission is *Cryptosporidium* [4]. Also research into *Blastocystis* has had a particular focus on genetic characterisation both in order to identify host specificity and to enable discovery of potentially pathogenic variants (see section 2.1). Various genes are used as markers; for CLIPPs, the SSU rRNA gene generally holds a lot of information, while for others, it is quite conserved across the genus, which is why other genes, typically house-keeping genes (*e.g.,* actin, elongation factor 1-alpha, heat-shock protein 70, *etc.*), or genus-specific genes (*e.g.,* glycoprotein 60 or beta-giardin for *Cryptosporidium* and *Giardia*, respectively) are chosen for better discrimination.

For CLIPP genera such as *Entamoeba* and *Blastocystis* molecular characterisation is relatively straightforward and has largely been based on specific PCR and subsequent Sanger sequencing. However, obtaining *Endolimax* and *Iodamoeba* SSU rDNA sequences can be challenging due to the typical absence of cultured material and the fact that SSU rDNA sequences of these parasites are relatively long (~2.5 kbp) [29, 30]. PCR using low-specificity eukaryotic primers preferentially amplifies any shorter and more abundant SSU rDNA from co-infecting/co-colonising organisms present in the intestine. This is often *Blastocystis* sp., which is frequently observed in *Endolimax*- and *Iodamoeba*-positive samples, as SSU rDNA sequences of *Blastocystis* are around 700 bp shorter than those of *Endolimax* and *Iodamoeba*. Even when specific amplification of *Iodamoeba*- or *Endolimax*-specific DNA is successful, Sanger sequencing of the PCR product will often result in a sequence of low quality (*i.e.,* with double peaks and sequence patterns lacking synchronisation) due to high intra-genome variation among the ribosomal gene copies, including differences in homopolymere length [31]. This makes Sanger sequencing of PCR products problematic and unable to clarify genetic diversity when used alone. Therefore, when relying on Sanger sequencing, a cloning step prior to sequencing has proven necessary for *Iodamoeba* and *Endolimax* SSU rDNA sequencing [29, 32]. This is probably the reason why accumulation of SSU rDNA sequences of these genera in the NCBI database has been relatively slow. Indeed, at the time of writing (January 2023), the number of *Iodamoeba*- and *Endolimax*-specific DNA sequences in GenBank are 79 and 34, respectively; these are quite modest numbers when compared to the number of sequences available for *Entamoeba* (~300,000) and *Blastocystis* (~75,000), and, for *Iodamoeba*, more than half of the sequences stem from one single study only and are practically identical.

As opposed to other areas within the field of microbiology, whole genome sequencing (WGS) is not yet a viable option for molecular characterisation of parasites. Very few parasites are readily established in culture and they are generally difficult to isolate in other ways. Moreover, whereas viruses and bacteria have relatively short genomes (*e.g.,* ~30 kb for severe acute respiratory syndrome coronavirus 2), parasites have genomes of typically about ~20 megabases (mb), which makes WGS expensive and time-consuming. And since CLIPPs and many other parasites are difficult to isolate from other organisms, the presence of competing template is typically an issue in amplification and sequencing procedures. Moreover, and not surprisingly given the situation described above, the reference data pipeline for WGS is still rudimentary, so even if WGS data would be available, genome data analyses would be challenging and time consuming. Nevertheless, given the diversity within CLIPPs already recognised, it would be fair to expect vast whole-genomic diversity, which could again be reflected in very diverse influences on the host and, maybe especially, host gut microbiota across CLIPPs; therefore, the mapping of genomes should receive priority.

Nevertheless, lately, the advantages of screening for genetic variation among CLIPPs using metabarcoding has become evident. Here, shorter fragments (a few hundred bp) of PCR-amplified SSU rRNA genes are sequenced using NGS technology, and consensus sequences can be generated manually from the total sequence output using DNA sequence clustering tools and become subject to genetic analysis, including

phylogenetic analysis. The method may even be sufficient for establishing hypotheses on new species/ribosomal lineages and can be used to demonstrate intra-genome variation in the genera *Iodamoeba* and *Endolimax*. However, near-complete SSU rDNA sequences should be obtained for potentially new lineage whenever possible for more robust phylogenetic inferences, and here, other technologies, such as MinION sequencing, are currently being tested.

During the period in which the publications included in the present thesis were published, research into genetic and host specificity of parasites went from relying mainly on traditional parasitological methods (including morphological observations) and conventional PCR and sequencing to relying increasingly on NGS-based approaches such as metabarcoding (or "meta-ribosomalomics"). This is reflected in the selection of publications included in this thesis, of which six involved the use of data obtained by metabarcoding. In the absence of morphological data, the applicability of the concept of cryptic species is challenged, and delineation of species would have to rely solely on phylogenetic species concepts. Nevertheless, metabarcoding has been of particular use in the mapping of parasites in wastewater samples, where it contributes to capturing of the diversity and relative prevalence of parasites in the society from which the wastewater is sourced [33].

1.5. Gut Microbiota Profiling

Despite trailblazing observations on microeukaryotic diversity in the human distal gut already in 2008 [34], until quite recently, high-throughput sequencing technologies used in studies of gut microbiomes have largely focussed merely on bacterial taxa. It may be so that the bacteria in the gut outnumber other members of the gut microbiome, namely archaea, fungi and non-fungal eukaryotes including CLIPPs; however, the latter may with their much larger genomes play roles that would be ignored, if studies relied on analysis of bacterial taxa only [35]. An advantage of the use of meta-ribosomalomics is the ability to compare bacterial communities in the digestive tract in individuals with and without parasites to test for significant differences in bacterial composition and to hypothesise to which extent such differences might be driven by parasite colonisation/infection.

Studying the distribution of parasitic genera common in humans in non-human hosts and the relationship between such parasites and bacterial communities in non-human hosts can help us hypothesise on the clinical and public health significance of these parasites in humans. Studies of wildlife mammals sampled in their natural habitats appear particularly interesting, since the gut microbiomes of these animals may largely be unaffected by factors reflecting human intervention and therefore would appear more original, potentially reminiscent of the original human gut microbiome given the evolutionary relationships between humans and larger mammals. Gut microbiota profiling of experimental animals before and after exposure to parasite

inoculation can also help 1) identifying factors that might drive the establishment of parasites in the intestine and 2) explore the impact of parasite establishment on gut ecology and host immune responses [36, 37].

1.6. Aim of Studies

The primary aim of the work reflected in this thesis have been to expand the knowledge on the genetic diversity and host specificity of CLIPPs found in the human digestive tract and mostly referred to as non-pathogenic, namely *Blastocystis*, *Dientamoeba*, *Endolimax, Entamoeba* and *Iodamoeba.* A secondary, and perhaps a more presumptuous aim was to try to inform public health policy makers and stake holders regarding the significance of single-celled parasitic genera commonly found in humans by obtaining a robust and reliable impression of the colonisation rate of some of these genera in both human and non-human hosts and by studying digestive tract microbiota profiles in association with the presence/absence of intestinal parasite colonisation.

2. GENETIC DIVERSITY & HOST SPECIFICITY OF CLIPPs OBSERVED IN HUMANS

In this section, there will be a brief account of the genetic diversity and host specificity of the most common CLIPPs found in humans, namely *Blastocystis*, *Dientamoeba*, *Endolimax*, *Entamoeba* and *Iodamoeba*.

Before that, I would like to highlight briefly a few methodological features that pertained to much of the work related to obtaining and analysing genomic data data in many of the studies:

- Genomic DNA was extracted from faecal/environmental samples or cyst preparations using typically the QIAamp DNA Stool Mini Kit (QIAGEN, Hilden, Germany) or the EasyMag or eMAG (NUCLISENS; bioMérieux, Inc.) protocols.

- Conventional PCRs used primers as appropriate, and Sanger sequencing was used to sequence PCR products. Two of the studies involved TA cloning [29, 32].

- Sequence alignments for the generation of consensus sequences and alignments for phylogenetic analyses were generated using a variety of free software; first and foremost different versions of MEGA [38-40], but also the online software Multalin (http://multalin.toulouse.inra.fr/multalin/) and Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/). In cases where consensus sequences were generated from sequence outputs from amplicon-based NGS, it could not be ruled out that a very minor fraction of the SNPs identified would reflect sequencing error rather than biological polymorphism. However, none of our analyses would be particularly sensitive to these errors, since these errors would either be diluted out by the vast variation seen between some subtypes (*e.g., E. gingivalis*; section 2.4.3), or they would be excluded from our multiple sequence alignments, as we would generally exclude bp positions where a SNP would be present in one sequence at a position where all other sequences were similar and if the sequence would otherwise be similar to another sequence in the alignment.

- For mitochondrion-like organelle (MLO) genome assembly [41], Staden Package [42] was used.

- Metabarcoding was used in some studies to identify positive samples and to screen for genetic diversity. Fasta files extracted from the BION server available at SSI were used to generate consensus sequences as appropriate [13, 33, 43-45].

- What is referred to as 'barcoding' [46-48] was used to identify subtypes of *Blastocystis* [47, 48].

- Phylogenetic analyses used distance-based analysis (Neighbor-Joining method), Maximum Likelihood, and Bayesian analysis as relevant [29, 32, 44, 49-53].

- Two articles included diversity analysis of gut or oral cavity bacterial microbiota in relationship with parasite colonisation, where the analyses included were alpha and beta diversity analyses and Linear Discriminant Analysis Effect Size (LEfSe) analysis [43, 45].

Small subunit rDNA sequences were sourced from the NCBI nucleotide database (https://www.ncbi.nlm.nih.gov/nucleotide/). Briefly, the species or genus name (as applicable and relevant) was entered in the search filed (in citation marks) along with Boolean operators as applicable (*e.g.,* "*Entamoeba coli*" AND "ribsosomal"), and the resulting list of hits was then subject to filtering using the filters and other features available at the site where relevant to ensure the inclusion of 18S sequences only and exported to Excel using the feature 'Send to', by choosing 'Complete record' + 'File' and the 'INSDSeqXML' option under 'Format'. The downloaded file would then be opened in Excel and sort functions used to identify relevant features, such as host information, sample material, date of data deposition, country of origin, *etc.*, where available. Excel pivot functions were used particularly for *Blastocystis*.

Near-exhaustive summaries of SSU rDNA gene sequences are provided for all the organisms, but in the thesis proper, only summaries for non-*Blastocystis* organisms were included, since the number of *Blastocystis* SSU rDNA sequences in the NCBI database is around 10,000 and therefore would take up too much space in this thesis, in my opinion. However, a full repository of Blastocytis SSU rDNA sequences is available at GitHub (https://github.com/Entamoeba/DMSc-Thesis/blob/main/Blastocystis%20SSUs%20from%20GenBank%20-%20Stensvold%2021012023.zip). The data downloaded from the NCBI database for the tables were edited only to a minimum (*e.g.,* spelling errors and information on host, where this had not been provided but was given elsewhere).

At least three issues generally hamper attempts to generate exhaustive species-specific SSU rDNA sequence repositories for CLIPPs – these are described and exemplified below:

1. If only sequences listed specifically as for instance '*Entamoeba coli*' or '*Entamoeba polecki*' are selected, then information from some of the many sequences deposited as '*Entamoeba* sp.' or 'uncultured *Entamoeba* clone' that are in fact '*Entamoeba coli*' or '*Entamoeba polecki*' is missed. Examples:
   a. KY658179 is listed as '*Entamoeba* sp.' but is '*Entamoeba coli* ST3'
   b. JX131936 and JX131943 are listed as 'Uncultured *Entamoeba* clone, but are in fact '*Entamoeba hartmanni*'

2. Sequences are seen listed as *e.g.,* '*Entamoeba polecki*' but cannot be verified as such (and maybe not even *Entamoeba*). Examples: AB845670, AB845671.
3. An *Entamoeba coli* sequence was observed listed as '*Entamoeba muris*' in GenBank (FN396613), and would therefore be missed, if only '*Entamoeba coli*' sequences were queried. There are also examples of situations where an *Entamoeba* has been misclassified as a completely different genus (*e.g.,* KU886548, which is listed as *Terfezia*, but which is *Entamoeba gingivalis*).

Another example of GenBank sequence entries where things have gone really wrong can be observed for the sequences MW133761–MW133772, which are all individually listed as '*Entamoeba* sp.' For this batch of sequences, only two are indeed *Entamoeba*, namely MW133765 and MW133768, which are both *Entamoeba coli*.

When assessing the validity of sequences potentially reflecting new intra-specific lineages, it may be helpful if already two or more lineages have been acknowledged within a species. If the variation observed in a new sequence is limited to parts of the gene where variation exists between the lineages already acknowledged, this could indicate that the sequence is valid (and not artefact). Meanwhile, if random SNPs are observed in places where the gene is otherwise conserved (typically between species of a genus), the sequence might be considered an artefact until confirmed. Sequence chimaeras (see section 2.3.6.), of which there are quite a few in GenBank, can distort phylogenetic analyses and should be recognised and removed from analyses.

Finally, it should maybe be mentioned that according to PubMed, the number of publications related to *Blastocystis* is 10 times the number of that of for instance *Iodamoeba*-associated publications, and therefore, the section on *Blastocystis* is more elaborate than sections on most of the other CLIPPs included.

2.1. BLASTOCYSTIS

2.1.1. Aspects of biology, life cycle, diagnosis, and epidemiology

The most common intestinal parasite in the human population might be a parasite that genetically does not resemble any of the other organisms parasitising on the human gut, namely *Blastocystis*. An oomycete sharing a most common ancestor with *Proteromonas* [54, 55], which is a flagellated single-celled parasite of the hind gut of lizards, it might colonise at least one billion people worldwide, and easily two, if data from recent DNA-based surveys of *Blastocystis* in European study populations can be extrapolated to the rest of the world. Indeed, in some populations, *Blastocystis* appears to be more or less an obligate finding in stool samples [56, 57]. *Blastocystis* is one of at least two Stramenopiles organisms known to parasitise on humans, the other one being *Pythium*, a cause of keratitis and endophthalmitis among other clinical manifestations [58].

Nevertheless, the life cycle and transmission of *Blastocystis* is quite reminiscent of that of many other CLIPPs, comprising a trophozoite-like stage (typically referred to as the 'vacuolar form') and a cyst stage [59] that makes *Blastocystis* amenable to faecal-oral transmission. There is some evidence that cysts might not always be detectable in *Blastocystis* carriers and that the detection of cysts in human faecal samples is independent of subtype and estimated gastrointestinal transit time [60]. Other stages reported include the granular and the amoeboid stages; however, it is not clear to which extent these stages reflect artefact stages. At least the granular stage may often be encountered in cultures and may reflect cells about to undergo apoptosis, as granular cells are often seen in cultures that have not been maintained, whereas rarely seen in fresh cultures (personal observations).

Indeed, *Blastocystis* is one of the few CLIPPs that can relatively easily be cultured [60-62], which expands the research opportunity for this parasite; however, *Blastocystis* is difficult to grow in the absence of bacteria in the culture medium. Cultures can be cryopreserved and thawed for future use [61]. With the specificity of culture being optimal [63], short-term *in vitro* culture as a diagnostic method appears superior to the traditional formol ethyl acetate concentration technique (FECT), but inferior to DNA-based detection [62]. Culture also appears to have high efficacy for *Blastocystis* in non-human hosts such as NHP and pigs [64, 65], although a more detailed view of the applicability to non-human samples of short term *in-vitro* culture as a diagnostic method is warranted. A positive culture result would evince colonisation and the presence of a live isolate in the host from which the sample is sourced, whereas a positive PCR result only shows that DNA of *Blastocystis* (whether dead or live) has been detected in the sample.

After developing several DNA-based methods for detection and subtyping [16, 62, 66], we developed a real-time PCR for more sensitive detection of *Blastocystis* [67], and this method was recently recommended for use in diagnostic laboratories [68]. Real-time PCR is possibly one of the most sensitive methods to distinguish

between carriers and non-carriers, and with faecal DNA already available, subtyping using state-of-the art methods [41, 47, 48] can be performed for research purposes. Recently, metabarcoding was proved to be a quite efficient way of detecting *Blastocystis* [13, 14], with the added benefit of enabling subtype determination based on the data output, including differentiation of multiple subtypes in any given sample [33].

Blastocystis is a rare example of an anaerobic eukaryote with organelles that have retained some mitochondrial characteristics, including a mitochondrial genome, and so, unlike most of the other CLIPPs included in this thesis, *Blastocystis* has not only one but two genomes: a nuclear genome of 12.9–18.8 mb (depending on ST) encoding 5713–6544 proteins, and a 'mitochondrial' genome of 27.7–29.3 kb [59]. The mitochondrial genome is that of the so-called mitochondrion-like organelle (MLO), and the first MLO genomes were published by Perez-Brocal and [69] and Wawrzyniak et al. [70] in 2008; these genomes represented subtypes 1 and 4 (for more information on subtypes, see section 2.1.2). A few years later, our group sequenced a few additional MLO genomes [41, 71] for two particular purposes; 1) to learn more about the genes contained in the MLO, and 2) to develop multi-locus sequence typing systems that might yield more genetic resolution than mere SSU rDNA sequence analysis.

Denoeud and colleagues published the first nuclear genome of *Blastocystis* in 2011 [72], which may very well be the smallest genome ever sequenced of a Stramenopiles organism, being just under 19 mb in size and containing about 6,000 genes.

Maybe one of the most interesting take-home message from studies of both MLOs and the nuclear genome is that *Blastocystis* might be able to tolerate environmental exposure that is not strictly anaerobic [73]. Still, *Blastocystis* appears to thrive in gut ecological niches that are dominated by anaerobes [45]. In humans, it is likely that *Blastocystis* may primarily be lodged in the terminal ileum, the caecum, and the proximal part of the colon.

The fact that *Blastocystis* is one of very few parasites found in humans that can easily be cultured [61, 62], makes it more amenable to genome-based studies than many other parasites. This is already reflected by the fact that there are several draft nuclear genomes in GenBank, some of which were deposited by our group, and there are also genomes representing the MLO for quite a few of the subtypes. Another potential opportunity of culture could be the induction of the cyst stage so that cysts could be produced for experimental studies; however, to my knowledge, a robust protocol for cyst induction is still to be developed.

Data on *Blastocystis* in the environment are scare. Maybe unsurprisingly, this parasite was an obligate finding in our study of Swedish wastewater samples [33]. In Thailand, a variety of subtypes were identified from *Blastocystis* cultured from water and soil samples [74, 75]. The parasite can survive exposure to chlorine and hydrogen peroxide [76].

In some human populations, *Blastocystis* is almost an obligate finding. All 93 children sampled in Senegal by El Safadi and colleagues tested positive [56], and in a study that we carried out on samples from 199 Nigerian children aged 2–14 years, we identified a positive rate of 84%, with prevalence increasing by age [57]. In children <4 years of age, *Blastocystis* was not nearly as common as in older children, and, in general, there was an almost linear relationship between colonisation rate and age. Since exposure to *Blastocystis* could be similar across all age groups, the association between age and *Blastocystis* colonisation independently corroborated by relatively high-powered studies from Brazil [77] and Libya [78] should be investigated in greater detail. It could be hypothesised that *Blastocystis* is dependent on a more mature (diverse) microbiota in order to be able to establish.

In our study of samples from Nigerian children, we found a clear inverse association between *Blastocystis* colonisation and malaria infection ($P < 0.0001$) [57]; however, children with malaria being younger than children with no malaria, this finding was attributed to the age effect of *Blastocystis* colonisation.
In *Denmark*, *Blastocystis* and *Dientamoeba* are by far the most common CLIPPs observed, but where *D. fragilis* is mostly seen in younger children, *Blastocystis* tends to colonise a wider range of age groups, being more common in adolescents and adults [79]. In adults, it is not uncommon to see both.

Blastocystis appear to be able to colonise the human intestinal tract for many years. Indeed, in our study led by Dr Scanlan, we observed that *Blastocystis* was present in a subset of healthy adult individuals sampled over a period of time between 6 and 10 years, indicating that it is capable of long-term host colonisation [80]. We based this assumption on the observations that the test individuals had tested positive for the same SSU rDNA sequence allele (see below) on samples taken several years apart.

The clinical significance of *Blastocystis* has been subject to a very heated debate over the past few decades. It has been speculated that *Blastocystis* could induce disease through elicitation of toxic-allergic reactions, degradation of human secretory immunoglobulin A by proteases, changes in epithelial permeability, induction of apoptosis of host intestinal cells and disruption of the epithelial barrier function, and/or modulation of immune response and cytokine release from colonic epithelial cells [81]. Some years back, we focussed on some of the limitations associated with pursuing the clinical significance of *Blastocystis* [81]. Case control, cohort, and randomized controlled clinical trials that could shed light on the role of *Blastocystis* in health and disease remain limited and are usually hampered by the use of methods of limited efficacy, including insensitive methods for evaluation of medical intervention. Moreover, *Blastocystis* is often observed together with other parasites, such as *D. fragilis* and species of *Entamoeba*, so investigating the isolated clinical effect of *Blastocystis* carriage is far from straightforward.

Nevertheless, we reviewed the efficiency of drugs used to eliminate *Blastocystis*, finding that the efficacy of metronidazole (MZ), which has been recommended [82] and used with an aim to ‘cure’ what has been called

'blastocystosis', was extremely limited [83, 84]. *In vitro*, MZ has high efficiency; however, we produced data indicating that herbal extracts of *Mallotus oppositifolius*, a medicinal plant traditionally used in sub-Saharan Africa to treat or alleviate stomach ache, diarrhoea/dysentery and diabetes, are almost just as efficient [85].

I was involved in a longitudinal, prospective case study, where 11 symptomatic patients positive for *Blastocystis* underwent outpatient clinical assessment to exclude other diagnoses before being treated for 14 days with either MZ 400 mg × 3/daily or trimethoprim (TMP)/sulfamethoxazole (SXT) 160/800 mg × 2 daily; none of the patients tested negative for *Blastocystis* following therapy [86].

I have moreover been involved in three cases for which data on the effect of treatment were available: One was the case of a woman with combined *Blastocystis* subtype 9 (ST9) and *D. fragilis* colonisation, who failed to experience clinical or microbiological effect of various treatments (**Table 3**) [84]. In another case, we were following a woman with ST8 as the only identified potential cause of intestinal symptoms, who achieved clinical and microbiological cure after completing 10 days of treatment with thrice-daily doses of 80 mg TMP/400 mg SXT; previous treatments with MZ had proven futile [87]. In the third case, a 20-year-old man with ST2 colonisation, who developed generalised urticarial, also failed to respond to MZ treatment and experienced clinical and microbiological cure only after adding paromomycin to the MZ treatment [88]. To my knowledge, no drug has to date proved 100% efficient in terms of eliminating *Blastocystis*.

Still, it could easily be argued that *Blastocystis* should not generally be included in routine diagnostic work-ups in clinical microbiology laboratories but rather be subject to scrutiny in research studies aiming to map the public health and gastroenterological impact of presence and absence of the parasite [89]. This conclusion is based on two major observations: Firstly, in Denmark, *Blastocystis* appears to be more common in the background population (positivity rate, 22%), while less common in patients with functional bowel disease (positivity rate, 15%) [90] and least common in patients with inflammatory bowel disease and acute diarrhoea (positivity rates, 0%–5%) [91, 92] (see section 4), which could suggest a protective role of *Blastocystis*. Secondly, since long, *Blastocystis* has by our own research groups (and since then also by many others) been hypothesized to be an indicator of a healthy gut microbiome and a normal body mass index (see section 4) [45, 93-96]. Thirdly, eradicating *Blastocystis* from the gut can be extremely difficult, and the treatments might do more harm than good, also in the long run. Finally, and very importantly, the current focus on antimicrobial stewardship should be prioritised. Overuse of for instance MZ can lead to antimicrobial resistance in both bacteria and parasites.

2.1.2. Genetic diversity and host specificity of *Blastocystis*

In 2006, we revisited the terminology of *Blastocystis*, which by then had been subject to a perplexing number of classification systems. In the process, we decided to abandon the species name '*Blastocystis hominis*' for reasons stated above (see Section 1.3), and consensus was also reached to adapt the subtype system still in place, and which is based primarily on the relative amount of diversity across *Blastocystis* SSU rRNA genes [26]. Since then, subtyping studies of *Blastocystis* from humans, other mammals and birds have led to a vast increase in publications on *Blastocystis*. This means that *Blastocystis* sequences from hosts such as amphibians, reptiles and insects are not included in the subtype terminology.

Before this time, molecular characterisation of *Blastocystis* had often been performed using the so-called sequence-tagged-site (STS) primers originally developed by Yoshikawa in 1998 [97]. These primers had been designed from random amplified polymorphic DNA sequences, with the nature of the DNA targets as well as their copy numbers remaining unknown. When I compared this method with barcoding published in 2006 by Scicluna and colleagues [46], I found that the latter method was more sensitive [48]. False-negative results by the STS assay were not linked exclusively to certain subtypes or alleles, and evidence of substantial genetic variation in STS loci was obtained. Over the next many years, the use of the STS method diminished, while the barcode method has been used by quite a few research teams to date. This has been advantageous for *Blastocystis* research, since standardisation of methods allows for inter-study comparisons, and DNA sequences obtained by the barcode method hold more information than just the subtype [41].

When the terminology was introduced, a total of nine subtypes were acknowledged [26]. Two years later, we published evidence of a new subtype, ST10, which we had identified in several Danish cows, but also in a sheep, a roe deer and in a lemur from a zoo [51]. Again a few years later, the number of subtypes increased to 17, as a few new subtypes were introduced: Our Australian colleagues identified ST11 from elephants, ST12 from giraffes, and ST13 from a quokka [98]; moreover, Fayer and colleagues introduced ST14 based on sequences obtained from cattle in the US [99]. In our large survey from 2013 going out from London School of Hygiene and Tropical Medicine and led by Dr Alfellani, we corroborated the validity of ST13, which had now been found in a mouse deer, and ST14, found in cattle and a mouflon [100]. Moreover, subtypes 15, 16, and 17 were introduced as new subtypes found in camel and gibbon (ST15), kangaroo (ST16; data deposited in GenBank by Hisao Yoshikawa), and a gundi (ST17).

Over the next 10 years, many more subtypes made it into the *Blastocystis* terminology, some of which were found invalid, and where we realised that the sequences that had been proposed as new subtypes represented sequence chimaeras [27]; these subtypes (ST18, ST19, ST20, and ST22) are now considered redundant. At the time of writing (January 2023), at least 34 subtypes have been acknowledged, and as the 'tree' has expanded,

it is becoming clear that the subtype collection might benefit from a re-evaluation to keep terminology practical (**Figure 2**). For instance, it might be useful to collapse ST24, ST25 with ST14 for two important reasons: 1) They cluster with very high bootstrap value and appear to exhibit less genetic diversity than what is observed within a subtype such as ST7; in fact, the two subtypes are not receiving individual support in **Figure 2**, which is based on 1,189 positions (and not near-complete genes, since these are not available for all subtypes), but are rather engulfed by ST14; 2) there is an overlap in host range. A somewhat similar situation is seen for ST21, ST26, ST30 and ST32, all of which are from ruminants, and which could be regarded as one subtype. However, any changes to the terminology might not be imminent, since it is perfectly operable for the time being, and it has not yet proved unuseful as such.

There are examples of sequences in GenBank that might represent new subtypes but for which additional sequencing and revisions of phylogenetic analyses are required. The sequences published in the study on *Blastocystis* in rabbits in China from 2022 by Su and colleagues is just one such example [101].

Efforts to identify whether 'new' sequences could represent a new subtype have been facilitated by sequencing technologies that were not in place until relatively recently. Especially MinION sequencing has proved particularly relevant, since it allows for cost-effective sequencing of near-complete SSU rRNA genes [102, 103], which was previously a laborious and difficult activity, but which was recommended to increase the quality of phylogenetic analyses [27, 73]. My own experience with MinION sequencing technology is limited, but we used it in a study of *Blastocystis* and *Entamoeba* in muskoxen (article in preparation).

There are now close to 12,000 nuclear SSU rDNA sequences (18S) for *Blastocystis* in the NCBI database (Supplementary Table 1 available on GitHub: https://github.com/Entamoeba/DMSc-Thesis/blob/main/Blastocystis%20SSUs%20from%20GenBank%20-%20Stensvold%2021012023.zip), and almost 10,000 for which host information was provided, so it is beyond the scope of this thesis to provide an exhaustive review of the host specificity of *Blastocystis*. The number of *Blastocystis*-associated publications has sky-rocketed over the past few years, with almost 750 being included in PubMed over the past five years. I will, however, try and draw up some features of *Blastocystis* host specificity in the following.

Our understanding of *Blastocystis* host specificity has been building up for now more than 15 years. Maybe more than 95% of colonised humans in Denmark and the rest of Europe have one or more of the subtypes 1, 2, 3, and 4 [80, 92, 104-108], and globally, ST3 appears to be the far most predominant subtype of the four in humans [109]. Interestingly, outside Europe, mainly subtypes 1, 2, 3 predominate, while ST4 is rare [14, 104, 110-113]. Other subtypes may rarely be seen in human hosts, and it is therefore likely that humans are not natural hosts of such other subtypes.

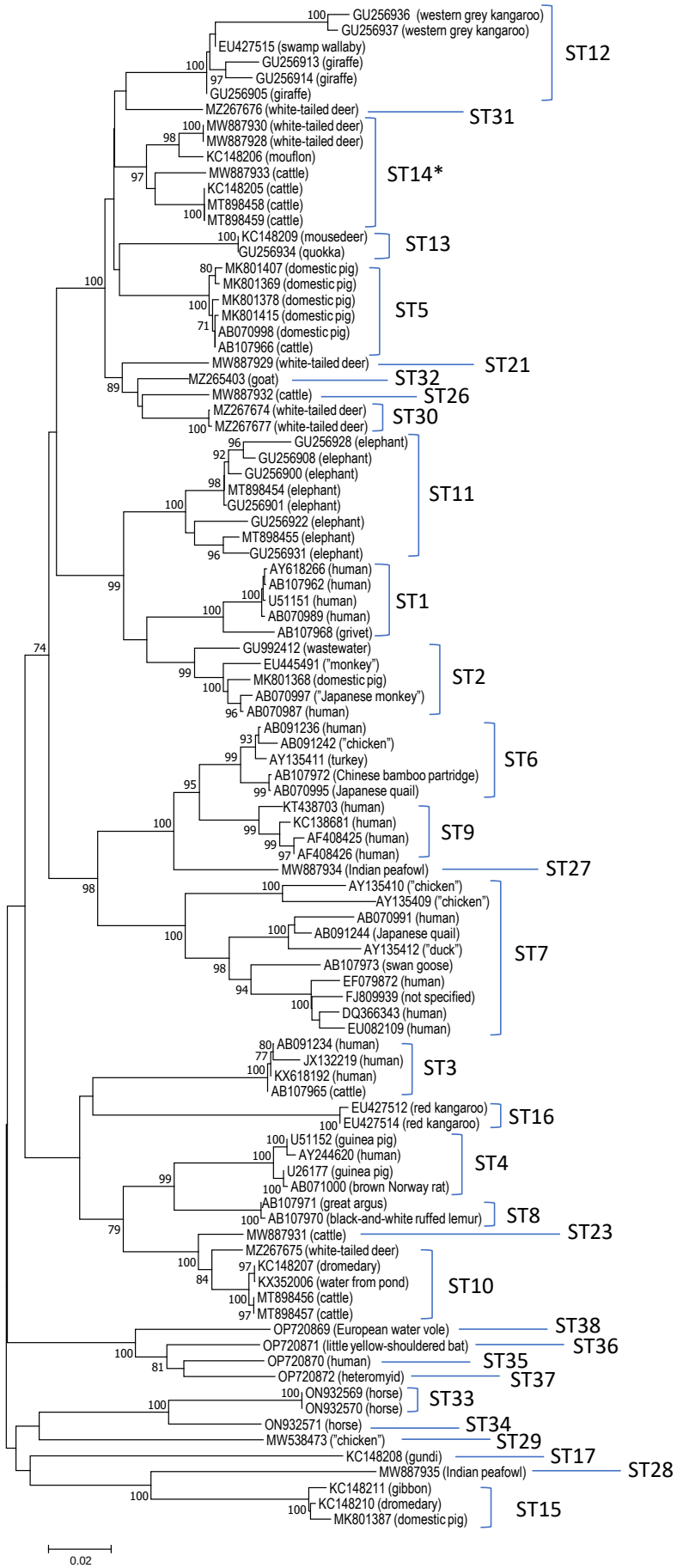

GU256936 (western grey kangaroo)
GU256937 (western grey kangaroo)
EU427515 (swamp wallaby)
GU256913 (giraffe)
GU256914 (giraffe)
GU256905 (giraffe)
ST12
MZ267676 (white-tailed deer)
ST31
MW887930 (white-tailed deer)
MW887928 (white-tailed deer)
KC148206 (mouflon)
MW887933 (cattle)
KC148205 (cattle)
MT898458 (cattle)
MT898459 (cattle)
ST14*
KC148209 (mousedeer)
GU256934 (quokka)
ST13
MK801407 (domestic pig)
MK801369 (domestic pig)
MK801378 (domestic pig)
MK801415 (domestic pig)
AB070998 (domestic pig)
AB107966 (cattle)
ST5
MW887929 (white-tailed deer)
ST21
MZ265403 (goat)
ST32
MW887932 (cattle)
ST26
MZ267674 (white-tailed deer)
MZ267677 (white-tailed deer)
ST30
GU256928 (elephant)
GU256908 (elephant)
GU256900 (elephant)
MT898454 (elephant)
GU256901 (elephant)
GU256922 (elephant)
MT898455 (elephant)
GU256931 (elephant)
ST11
AY618266 (human)
AB107962 (human)
U51151 (human)
AB070989 (human)
AB107968 (grivet)
ST1
GU992412 (wastewater)
EU445491 ("monkey")
MK801368 (domestic pig)
AB070997 ("Japanese monkey")
AB070987 (human)
ST2
AB091236 (human)
AB091242 ("chicken")
AY135411 (turkey)
AB107972 (Chinese bamboo partridge)
AB070995 (Japanese quail)
ST6
KT438703 (human)
KC138681 (human)
AF408425 (human)
AF408426 (human)
ST9
MW887934 (Indian peafowl)
ST27
AY135410 ("chicken")
AY135409 ("chicken")
AB070991 (human)
AB091244 (Japanese quail)
AY135412 ("duck")
AB107973 (swan goose)
EF079872 (human)
FJ809939 (not specified)
DQ366343 (human)
EU082109 (human)
ST7
AB091234 (human)
JX132219 (human)
KX618192 (human)
AB107965 (cattle)
ST3
EU427512 (red kangaroo)
EU427514 (red kangaroo)
ST16
U51152 (guinea pig)
AY244620 (human)
U26177 (guinea pig)
AB071000 (brown Norway rat)
ST4
AB107971 (great argus)
AB107970 (black-and-white ruffed lemur)
ST8
MW887931 (cattle)
ST23
MZ267675 (white-tailed deer)
KC148207 (dromedary)
KX352006 (water from pond)
MT898456 (cattle)
MT898457 (cattle)
ST10
OP720869 (European water vole)
ST38
OP720871 (little yellow-shouldered bat)
ST36
OP720870 (human)
ST35
OP720872 (heteromyid)
ST37
ON932569 (horse)
ON932570 (horse)
ST33
ON932571 (horse)
ST34
MW538473 ("chicken")
ST29
KC148208 (gundi)
ST17
MW887935 (Indian peafowl)
ST28
KC148211 (gibbon)
KC148210 (dromedary)
MK801387 (domestic pig)
ST15
0.02

Figure 2 (previous page). Phylogenetic relationship between the subtypes of *Blastocystis* acknowledged to date. For higher resolution, please visit file repository on GitHub (https://github.com/Entamoeba/DMSc-Thesis/blob/main/Blastocystis%20Tree%20GitHub.pdf). The evolutionary history was inferred using the Neighbor-Joining method (Kimura 2-parameter method). The analysis involved 96 nt sequences. There were a total of 1,189 positions in the final dataset. The scale bar indicates nt substitutions per site.

Looking at non-human primates sampled in the old world, including both apes and monkeys, these tend to harbour the same subtypes as seen in humans, except for ST4. Meanwhile, we have published data suggesting that ST5 might be a common finding in apes, although potentially not in monkeys [114]; however, we found a few different species of cercopithecines in Thailand colonised by ST5 [65]. The predominance of ST1–ST3 in non-human primates [114], which happen to be the STs most commonly associated with humans, might suggest that these STs have a shared co-evolutionary history with humans and their closest living relatives. ST4, by contrast, is uncommon in non-human primates, whereas it is common in humans in Europe; contrary to ST1–ST3, ST4 from humans appears genetically conserved (see below), indicating a recent entry in the human population [115].

There is a fair chance that ST3 is the most common subtype in both human and non-human primates, and one might easily jump to the conclusion that NHPs constitute a reservoir for human ST3 colonisation; however, based on MLO genomic analysis, we produced evidence of cryptic genetic diversity within ST3, suggesting host-adapted genotypes (*i.e.,* variants within a subtype) of ST3, something that we confirmed by allele analysis of ST3 found in cercopithecines sampled in Thailand [65]. The cryptic diversity disclosed by studying MLO genome data was to a high degree reflected in ST3 SSU rDNA sequences, which is why the 'allele' concept and the publically available MLST database (https://pubmlst.org/organisms/blastocystis-spp) was introduced [41]. The allele system takes advantage of any intra-ST variation detected in *Blastocystis*. Sequences that differ down to one SNP can be submitted to the MLST database and is assigned a unique allele number. These allele numbers can then be used in studies calling for more subtle discrimination of strains than can be provided merely by providing information on subtype.

If we look specifically at arboreal monkeys, we found ST8 in monkeys of Asian or South American origin [114], but apart from these, ST8 has only relatively rarely been reported in primates.

From the research undertaken so far, the extent to which livestock and other synanthropic animals contribute to transmission of *Blastocystis* to humans remains somewhat unclear. A preponderance of ST10 and ST14 is seen in ruminant hosts sampled across the world, neither of which are often observed in humans. Meanwhile, the subtypes typically seen in suid hosts are subtypes 1, 3, 5, and 15 [13, 116], and even ST2 has been found in pigs [116]. Of these, both ST1, ST2 and ST3 are commonly seen in humans, but we currently do not know if ST1, ST2 and ST3 strains from pigs are different to ST1, ST2 and ST3 strains colonising humans, as the

allele system may not hold sufficient discrimination to tell these apart [13]. Analysis of complete SSU rDNAs and/or MLO genome markers might enable researchers to provide answers to this question.

Avian subtypes, such as ST6 and ST7 may occasionally be seen in human faeces [117], and ST8 is also occasionally seen. We also found ST8 in quite a few samples in our study of Swedish wastewater [33]. Since ST8 is uncommon in humans, and since the many other parasites found in the wastewater material were parasites known to colonise humans, we speculated that ST8 might stem from an animal that can live in in sewers, for instance rats, and that rats might even be natural hosts of ST8. Galán-Puchades and colleagues identified a *Blastocystis* colonisation rate of 83.5% in rats sampled in parks and sewers of Barcelona [118]; however, no subtyping of these samples was performed in the study. ST4, which is the sister taxon of ST8 (actually, ST8 was split out from ST4 when we established the consensus terminology), has been shown to colonise rats and other rodents easily [119-122] (see below), so the hypothesis might not be farfetched. More extensive sampling of rats is necessary to complete this picture. Of note, ST1 of human origin can be established experimentally in rats, with colonisation lasting for more than a year [37].

Until very recently, ST9 appeared to be an extremely rare finding and potentially restricted to humans hosts [84]; however, in 2021, Liu and colleagues published extensive evidence of ST9 in peafowl in China, and poultry was also identified as hosts of ST9 in Malaysia (GenBank entry KX234596). Indeed, it would not be surprising if birds were confirmed natural hosts of ST9, since this subtype is a sister taxon to ST6 (ST9 was split out from ST6 in when we established the consensus terminology), and since also ST27, which is also observed in birds, clusters together with ST6, ST7 and ST9 (**Figure 2**).

Many subtypes have more or less exclusively been detected in ruminants or at least in large, mainly herbivorous or omnivorous mammals, including marsupials. These include the subtypes found in the upper third of the tree (**Figure 2),** which is the part of the tree that I would refer to as the ST5-ST14 clade, and which includes the subtypes 5, 12, 13, 14, 21, 24, 25, 26, 30, 31, and 32. An interesting feature of this clade is that white-tailed deer have been found to host quite a few of the subtypes in this clade, including ST14, ST21, ST30 and ST31. Although much more limited in the number of subtypes, there is another artiodactyle-specific clade made up by ST10 and ST23, which again includes white-tailed deer as host among cattle and dromedaries.

The base of the phylogenetic tree has been expanded a bit recently by a number of new subtypes that tend to sit on relatively long branches, which indicates a substantial degree of genetic diversity. Some of these have been found in hosts that might not previously have received much attention in *Blastocystis* research, such as bats and heteromyids.

Overall, the most clear host specificity is seen mainly for the ST5-ST14 clade, comprising 'artiodactyle subtypes', which could be considered an analogue of the 'bovis complex' of *Entamoeba* (see section 2.3.) and for the clade made up by the avian subtypes (mainly ST6 and ST7). Apart from this, the lack of clear co-evolutionary trends could indicate that *Blastocystis* has been entering most major host groups more than once.

When having a rough look at the breakdown of hosts for the >9,800 SSU rDNA sequences in GenBank for which host information is available, humans account for a good half of these (56%). Cattle contribute at least 10% of all sequences, and suids 7%. Non-human primates account for at least 7.5% of the sequences, avian hosts 6%, and goats and sheep together 3% of the sequences. Larger mammals like elephants, giraffes, lamas, pandas, horses and deer contribute about 5%-6%, and rodents 1%. Other sequences are from lagomorphs, marsupials, reptiles, and even fish (*Clupea harengus* and *Pollachius virens*) and cockroaches. Strictly carnivorous mammals and omnivorous scavenger animals, on the other hand are relatively scarcely represented, with about 34 sequences from dogs (0.3%), 22 from cats (0.2%), 19 from foxes (0.2%), 11 (0.1%) sequences from bears, and 8 (<0.1%) from raccoon dogs. To this end, carnivores do not appear to host any particular subtype but rather a large variety, including subtypes 1-4, ST7, ST10, and ST14, to mention some, which together with the small number of sequences and the possibly very low prevalence of *Blastocystis* could indicate that carnivores may not be natural hosts of *Blastocystis* (also see Section 5). A study adding support to this hypothesis is the one by Heitlinger and colleagues, who characterised the eukaryotic and bacterial faecal microbiota of 42 spotted hyenas, and who did not report any finding of *Blastocystis* [123].

We (and many other teams) have highlighted the apparent absence or rarity of ST4 in human populations outside Europe [14, 57, 111, 124-129]. Although rare in humans outside Europe, this subtype has meanwhile been an occasional finding in non-human hosts. In a metanalysis of data published from the Americas, Jiménez and colleagues identified the observation of ST4 in 1.7% of humans sampled and in 7.2% of samples from non-human hosts [130]. Extending the metanalysis approach to global scale, Barati and colleagues [120] identified 18 studies (16 of which were studies of rodents sampled outside of Europe) that had used molecular methods to detect and differentiate *Blastocystis* subtypes in rodent samples, and ST4 was the most common among these.

Genetic variation exists within ST4, and already in 2011, we could differentiate at least two major clades; one that appeared common in both human and rodent hosts and one that we thought might be limited exclusively to rodent hosts [92]. However, the year after, we had produced evidence of a ST4 sequence from a human clustering in the clade thought to be strictly rodent (JN682513) (**Figure 3**). There are still two major clades of ST4, of which one (Clade 2) appears to be 'predominantly rodent', while the other one (Clade 1) contains sequences from both rodents, humans, and other mammals, a finding more recently corroborated by Katsumata and colleagues [119]. An interesting feature of these two clades is that where Clade 1 is practically clonal (*i.e.,* with very limited genetic variation), Clade 2 exhibits genetic heterogeneity. It could be speculated

that the introduction and expansion of Clade 1 in humans happened once, in Europe and relatively recently. Analysis of MLO genomes of ST4-positive faecal DNAs representing both clades confirmed the genetic homogeneity of Clade 1 [41]. Very interesting is the fact that Betts and colleagues identified *Blastocystis* in all 38 water voles sampled in their study, observing a clear predominance of ST4 [131], and there were examples of both clades of ST4 among the sequences.

Efforts to disclose whether some *Blastocystis* subtypes are more commonly seen in individuals with GI symptoms than in gut-healthy individuals have not produced any major breakthroughs. In one of our studies, it appeared that those harbouring zoonotic subtypes such as ST6, ST7, and ST8 might be more prone to experiencing symptoms than those who did not harbour these subtypes [117]. To this end, we published a case of ST8, where clinical and microbiological resolution was reached after treatment with TMP-STX [87].

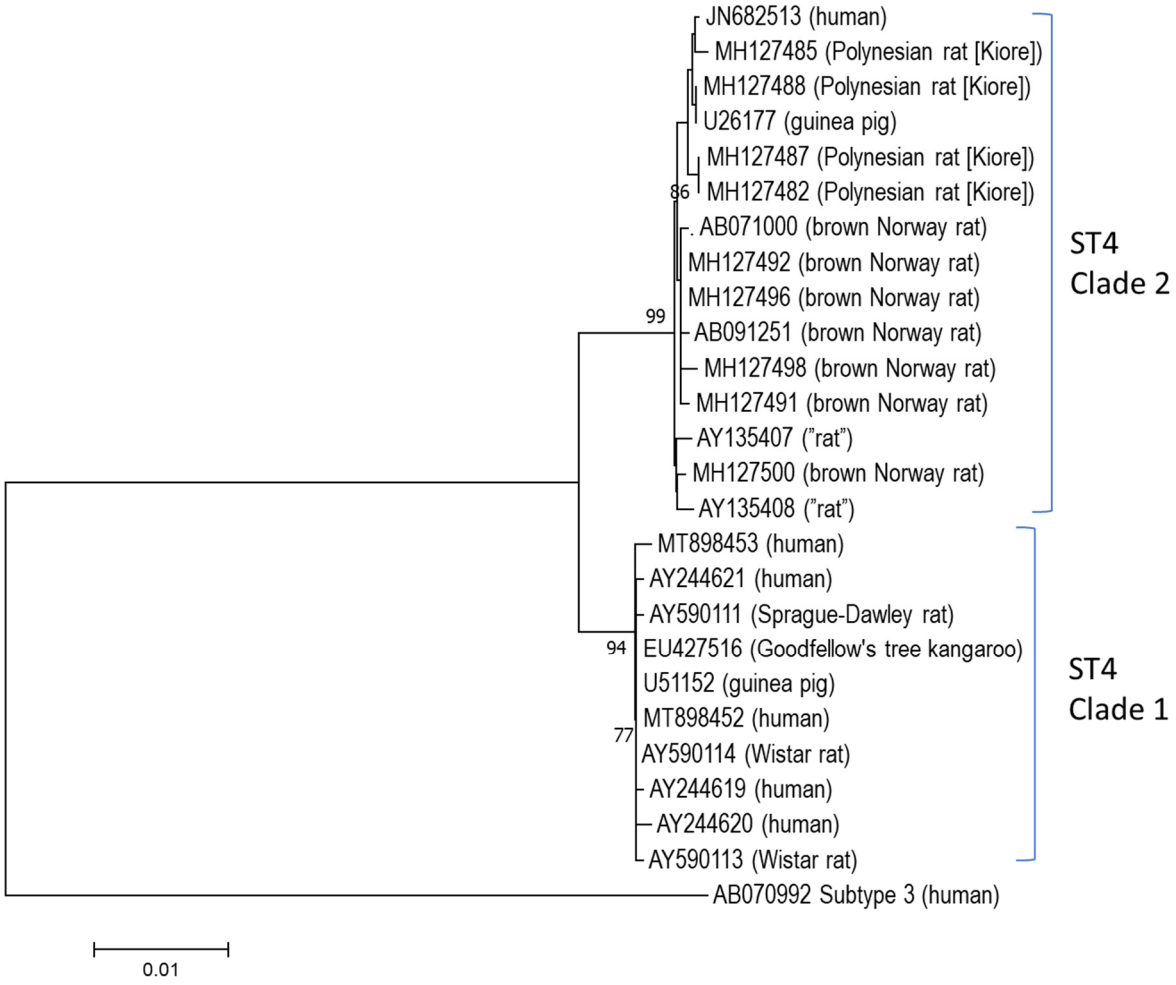


Figure 3. Analysis of intra-subtype diversity of *Blastocystis* ST4 SSU rDNAs. Subtype 3 was used as outgroup. There were a total of 1,734 positions in the final dataset. Two clades are appreciated, one of which (Clade 1) is genetically homogenous and holds sequences from several host groups, and the other of which (Clade 2) is genetically heterogeneous and holds sequences almost exclusively from rodents. The scale bar indicates nt substitutions per site.

In one of our other studies, we found almost exclusively ST4 in patients with acute diarrhoea sampled in Denmark [92], which at first glance could suggest a link between ST4 and this clinical condition; however, the positivity rate in the study was low compared with that of other groups of individuals studied in Denmark, and instead of reflecting an association between ST4 and diarrhoea, this observation may indicate that ST4 might be more resilient to any gut microbiota disturbances experienced during periods of diarrhoea than other subtypes. It should be noted that all the ST4 sequences identified in that study clustered in Clade 1 (data shown in the article).

To this end it should be mentioned that the use of traditional PCR followed by Sanger sequencing, which is the backbone of the barcoding method [46-48] and which many teams have used to characterise *Blastocystis*, may have led to an underestimation of mixed ST colonisation. Indeed, our use of metabarcoding has disclosed that mixed subtype colonisation appears to be a quite common phenomenon [14], and we recently concluded that the combination of real-time PCR with metabarcoding would be beneficial for epidemiological and surveillance studies [108, 132]. We also prioritised the development of subtype-specific primers to map the extent of mixed subtype colonisation [133], which is an approach that can be taken by those who do not have access to 'omics' technologies.

Over the past few years, we and others have produced data indicating strong links between *Blastocystis* and gut microbiota signatures and a role for *Blastocystis* as an indicator of gastrointestinal health. This is dealt with in more detail in section 4.

2.2. DIENTAMOEBA

2.2.1. Aspects of biology, life cycle, diagnosis, and epidemiology

Although described by Jepps and Dobell as early as in 1918 [134] and although being an extremely common coloniser of humans, the number of articles listed in PubMed rendered when searching for "*Dientamoeba*" is only about 429 at the time of writing (January 2023), which is about five times less the number of articles available on *Blastocystis*.

Despite its name, *Dientamoeba* is not an amoeba, but a flagellate that lost its flagellum [135]. Only one species, *D. fragilis*, is known. It is genetically related to *Giardia*, and shares a most common ancestor with *Histomonas meleagridis*, a protozoon that causes 'histomoniasis' in poultry. The observation of a cyst stage was reported relatively recently [136] but remains to be confirmed by additional teams. Indeed, microscopy of faecal concentrates from *D. fragilis* carriers would very likely reveal cysts of *Dientamoeba*, should these exist; the question remains as to whether anyone would recognise these. With regards to the sister taxon *H. meleagridis*, no cyst stage has been identified. Conspicuously, rather than surviving outside the host on cyst from, it would appear that *H. meleagridis* is transmitted by *Heterakis gallinarum*, a nematode of poultry [137]. Given the apparent lack of a cyst stage for *Dientamoeba* and given the genetic similarity to *H. meleagridis*, it is hypothesized that this parasite takes advantage of a transmission mode similar to that of *H. meleagridis*. A few years back, we found evidence of DNA of *D. fragilis* inside eggs of *Enterobius vermicularis* that we had surface sterilised with hypochlorite in order to remove the risk of contamination [138]. Although this does not prove that *D. fragilis* can be transmitted by pinworm, it suggests that further investigation of pinworm as vehicle for *D. fragilis* would appear relevant. In a registry-based retrospective cohort study of 9,945 patients tested for *D. fragilis* at the SSI between 2008 and 2011 we identified that mebendazole (MB) exposure was associated with increased risk of testing positive for *D. fragilis* [139], and since MB is practically only used to treat pinworm infections in Denmark, this could be interpreted as indirect evidence of an association between pinworm and *D. fragilis*.

We published the first data on *D. fragilis* in Denmark in 2007 [140]. We found a positive rate of 12% among patients with suspected enteroparasitic disease in the Copenhagen metropolitan area. Since the international scientific literature at that time suggested that *D. fragilis* should be considered an intestinal pathogen in humans [141, 142], routine testing for *D. fragilis* was initiated as part of the general parasitological workup for patients with diarrhoea or other intestinal symptoms, such as abdominal pain. The diagnosis of *D. fragilis* typically relies on examination of permanently stained fixed faecal smears or DNA-based detection using for instance real-time PCR. In 2007, Verweij *et al.* published a real-time PCR assay for specific detection of *D. fragilis* [143]; this assay was implemented in the Laboratory of Parasitology, SSI shortly after.

Six years later, we reviewed data from the routine testing for *D. fragilis* at SSI and identified a positivity rate of 43% across the more than 22,000 samples tested over a four-year period [144], with the positivity rate peaking at 71% in children aged ~7 years.

After i) dismissing any statistically significant microbiological or clinical effects of MZ treatment of children with abdominal pain and *D. fragilis* colonisation [145], ii) demonstrating that practically all children in institution in Denmark would be or become colonised by *D. fragilis* over a period of a year [146], and iii) showing that the organism was more common in the background population than in patients with GI symptoms [90], it was decided to exclude *D. fragilis* testing from the routine parasitological workup panel at SSI. This is a clear example showing the importance of having the level of parastisim in the background population inform decisions on which parasites to test for in the CML.

Indeed, we investigated the presence of *D. fragilis*-specific DNA in faecal samples obtained from a cohort of 142 0–6-year-old children that we had already tested for a diverse range of gastrointestinal pathogens [147]. Among the 108 children who had submitted two or more samples and thereby included in a longitudinal analysis, 32 tested *D. fragilis*-negative on the first sample but positive later, and the last sample from each of the 108 children was positive [146]. Risk factors associated with testing *D. fragilis*-positive including being >3 years old and having a history of recent travel abroad.

Colonisation by *D. fragilis* appears to be very common in Denmark and other European countries. However, in countries far away from this region, such as Australia where comparable diagnostic tools are used for detection and where there may not be large differences in testing strategies or populations tested (*e.g.,* in terms of age), the colonisation rate appears to be significantly lower (**Table 4**). It is therefore possible that the clinical and public health significance of *D. fragilis* colonisation differs according to geographical region. Indeed, it should be investigated whether the overall prevalence of pinworm infections mirrors the overall prevalence of *D. fragilis* colonisation; for instance, it may be so that the prevalence of pinworm in Australia is much lower than that seen in Europe; however, data on this remain scarce.

In a study aiming to investigate whether symptom relief could be obtained in *D. fragilis*-positive individuals with IBS using either MZ or tetracycline, we noticed microbiological responses in 15 of 25 individuals (60%), all by MZ; a clinical response was observed in 7 of 22 patients (32%), all by MZ. Meanwhile, some test individuals were insufficiently treated by MZ [148].
We were involved in a patient case where GI symptom relief was achieved upon eradication of *D. fragilis*, which proved successful using PM but not MZ [149]. In another case already mentioned (**Table 3**), *D. fragilis* was eradicated only after administration of PM 500 mg + MZ 750 mg thrice daily/10 days.

A cluster of *D. fragilis* infections associated with peripheral eosinophilia (PE) in a family that had been sharing a meal of shashimi (raw fish) was presented by Gray and colleagues [150]. To my knowledge, *Cystoisopora* is the only gut protozoon that can trigger eosinophilia [151], an immune response typically induced by helminth infections and allergic reactions. We thought that the PE might reflect to exposure to live anisakids that might have been present in the fish consumed by the family [152].

2.2.2. Genetic diversity in and host specificity of *D. fragilis*

What was not done in the case report by Halkjaer et al. [149] mentioned above was typing of the *D. fragilis* strain identified. *D. fragilis* comprises two genotypes, 1 and 2, of which genotype 1 is by far the most prominent; not only in Denmark, but also in other countries [153-157].

In 2000, Johnson and Clark demonstrated cryptic genetic diversity in *D. fragilis* by applying restriction-fragment length polymorphism (RFLP) analysis to SSU rDNA amplicons, establishing the existence of the two genotypes [158]. A few years later, Peek and colleagues used both PCR and RFLP and Sanger sequencing to screen *D. fragilis*-positive individuals sampled in the Netherlands for genotypes (referred to as "haplotypes" in that study), finding only genotype 1 [159]. Similar work was performed by Stark and colleagues and our own group a couple of years later [153, 160] in Australia and Denmark, respectively. In the study by Stark and colleagues, only genotype 1 was identified, and in our own laboratory, we also only identified genotype 1 (unpublished observations[1]). It is not known whether these two genotypes differ in terms of clinical significance. The genetic differences between the two genotypes across the entire SSU rRNA gene is ~3.5%.

Windsor and colleagues attempted sequencing of the Internal Transcribed Spacer (ITS) region of *D. fragilis*, but concluded that the value of this approach was limited due to intra-strain genetic heterogeneity [161]. 'C-profiling' was developed as a means of extracting useful data from sequenced ITS clones [162], but, to the author's knowledge, this method has only been used once ever since and on a very limited material [163], and therefore, the applicability and epidemiological relevance of this method remains uncertain.

In 2012, we noticed that two *D. fragilis* genotype 1 housekeeping genes had been amplified and sequenced, namely actin and elongation factor 1 alpha, both of which sequences were present in the NCBI database. Based on these sequences, we developed specific primers for PCR and sequencing and screened *D. fragilis*-

[1] In the abstract published in the Proceedings from the 5th European Congress on Tropical Medicine and International Health 24–28 May 2007 Amsterdam, the Netherlands by Stensvold et al. [153], there were no data from lab analyses (the work was ongoing at the time). However, the results are available in the MSc thesis by co-author Kenneth Dinesen that can be made available upon request.

positive genomic DNAs extracted from stool for genetic variation across these two loci plus the SSU rRNA gene. Our data indicated that genetic analysis of these three *D. fragilis* housekeeping genes enabled clear distinction between the two known genotypes, and phylogenetic analysis of translated, concatenated sequences confirmed the phylogenetic position of *D. fragilis*. Meanwhile, integration of housekeeping genes in multi-locus sequencing tools for *D. fragilis* would possibly have limited epidemiological and clinical value due to no further added genetic resolution [52]. In a multi-centre study a few years later, this conclusion was further corroborated [157], when we applied six new genetic markers to *D. fragilis*-positive samples from individuals from Italy, Denmark, Australia, and Brazil. Here, only one of 111 samples was positive for genotype 2, while the rest exhibited the genotype 1 profile; importantly, no further genetic resolution could be identified as potential markers for further discrimination.

Since then, the genetic diversity within *D. fragilis* has been studied only to a limited extent. David and colleagues identified only genotype 1 in a study carried out in Brazil [164]. Oliveira-Arbex and colleagues identified a *D. fragilis* positivity rate of 10.3% among 156 asymptomatic children in day care centres in São Paulo State, Brazil. The 16 positive samples were typed, identifying 14 genotype 1 and two genotype 2 sequences [156].

The data published to date bear witness of a low level of polymorphism and, as also mentioned by Caccio [165], are compatible with a clonal population structure of *D. fragilis*, which is in stark contrast to for instance *Blastocystis*. For *H. meleagridis*, the sister taxon of *D. fragilis*, two genotypes have been identified, and also here, genotype 1 appears to be predominating, although with some intra-genotype genetic variation [166].

To date, there are at least 52 SSU rDNA sequences (18S) available (**Table 5**) + quite a few sequences from the study by Windsor and colleagues [161], who sequenced the part of the ribosomal operon containing the ITS regions 1 and 2.

Since genotype 2 appears to be rare in studies involving human faecal samples, it could be speculated that this genotype would have a reservoir in one or more non-human hosts, and screening faecal samples from non-human hosts for *D. fragilis* therefore appears relevant.

The number of studies that involved testing of faecal material from non-human hosts is still limited, and an overview of these studies is provided in **Table 6**. It should be noted that there is a handful of reports on findings of *D. fragilis* in rats and non-human primates [167-170]; however, as these data are based on morphology only, we have chosen not to include the data in **Table 6**. At the time of writing (January 2023), there are *D. fragilis*-specific DNA sequences only from humans, pigs and budgerigars in the NCBI database.

As seen in **Table 6**, only genotype 1 has been identified in studies of non-human hosts that involved genotyping. Hence, any reservoir for genotype 2 remains to be identified. The situation is somewhat similar to that of the microsporidium *Enterocytozoon bieneusi* genotype C, which has been linked to two conspicuous outbreaks in Scandinavia [2, 6] and observed in immunosuppressed individuals [171, 172], but which has otherwise only rarely been reported of in humans, and which has only been detected in a couple of instances in non-human hosts (*e.g.,* in two gorilla samples in Rwanda [173] and a handful of mice in Spain and in central Europe [174, 175]), despite quite extensive screening of various animals. However, *E. bieneusi* genotype C was identified in quite a few of the Swedish wastewater samples [33] tested for Amoebozoa and *Blastocystis* (unpublished data). If more DNA were available, the Swedish wastewater samples could be screened for *D. fragilis* genotype 2 with specific primers.

In the event that *D. fragilis* is indeed transmitted by pinworm, the question arises which transport organisms that might be used by *D. fragilis* when colonising non-human hosts, which are not natural hosts of pinworm. It may be so that other nematodes might serve as vectors for *D. fragilis.*

2.3. ENTAMOEBA

Entamoeba is a genus comprising aflagellated endobiotic taxa, most of which parasitise the guts of vertebrates. Within the Archamoebae, *Entamoeba* forms a sister taxon to the genus *Pelomyxa*, which comprises large, free-living, flagellate, multinucleate amoebae [176] that can be found in anaerobic or microaerobic bottom sediments of stagnant freshwater ponds or slow-moving streams. As a result of adaptation to the anaerobic gut environment, Entamoebas do not have mitochondria, but mitosomes, which are considered mitochondrial remnants in a way similar to what is seen for *Blastocystis*, in which they are called MLOs, but which lack a detectable organellar genome [177].

Entamoebas are motile, can phagocytose bacteria, and, at least in humans, they typically colonise the parts of the gastrointestinal tract that have most bacteria; *i.e.,* the oral cavity and the colon. A couple of species have invasive properties and have been linked to disease in humans and non-human hosts (*E. histolytica*, *Entamoeba nuttalli*, and *Entamoeba invadens*), although asymptomatic carriage may be common. A couple may be found primarily or only in the environment (*e.g., Entamoeba marina, Entamoeba moshkovskii*) (**Table 7**).

Morphologically, organisms belonging to the genus *Entamoeba* are easily recognised by their nuclear features. Parasitologists will be familiar with the "ring-and-dot" appearance of one to several nuclei that can be observed in cysts. Most Entamoebas produce cysts, an exception being *E. gingivalis*. Mature cysts typically have either one, four, or eight nuclei, depending on species (**Table 7**). Each nucleus has a karyosome, which may be centrally located (*e.g., E. histolytica*) or eccentric (*e.g., E. coli*), and peripheral chromatin is present, which may be distributed in a fine, homogenous distribution (*e.g., E. histolytica*), or in lumps (*e.g., E. coli*).

While it may be straightforward to identify an organism as belonging to the genus of '*Entamoeba*', the differentiation of species within *Entamoeba* is associated with a lot more difficulty. Firstly, humans can host several established species of *Entamoeba* (**Table 7**), some of which are species complexes, and some of which cannot be differentiated based on cyst morphology. For instance, *E. histolytica*, *E. dispar*, *E. moshkovskii*, *E. nuttalli,* and *E. bangladeshi* all produce cysts of similar size and with four nuclei. *E. nuttalli* has only been reported once in a human [178], and there are still only scarce reports of *E. bangladeshi* [179, 180]. Since *E. nuttalli* is and *E. bangladeshi* might be pathogenic (the clinical significance of *E. bangladeshi* remains unresolved and there are only six SSU rDNA sequences in GenBank), these two species will not be subject to separate discussions in this thesis. Secondly, surveys using only morphological identification have been challenged by the fact that cysts of each species of *Entamoeba* might exhibit a continuum of morphological features, especially features such as cyst size and number of nuclei, depending on the maturity of the cysts.

Given this situation, it may not be a surprise that *Entamoeba* taxonomy has been subject to two types of "error" in the naming of species, which also pertain to the situation for other CLIPPs [181]: (1) reliance on a character that does not reflect underlying genetic divergence, leading to overestimation of diversity and the naming of invalid species: examples include relying on the host as a species-specific character when in fact some *Entamoeba* species have quite a broad host range; and (2) perceived morphological simplicity means that genetic divergence is not always reflected in morphological differences, which leads to underestimation of diversity and assigning the same species name to quite different organisms, a situation that we refer to as 'cryptic genetic diversity' (see below). To this end, and as a clear example, both of these considerations also led to dismissal of the species name '*Blastocystis hominis*' back in 2007 [26].

Nevertheless, over the past 20 years, the genetic universe of *Entamoeba* has been unfolding bit by bit (**Table 7**, **Figure 4**), and seven of the works included in this thesis [13, 33, 43, 44, 49, 50, 53] have contributed to what today is known about genetic diversity in *Entamoeba* hosted by humans and other larger mammals, such as cattle and pigs, and in waste water. State-of-the-art terminology of *Entamoeba* involves the use of 'species', 'subtype', 'ribosomal lineage' and 'conditional lineage', depending on the amount of genetic diversity and the type of information available (sequence length, morphology, host, *etc.*; **Table 2**, **Table 7** and

Table 8).

The use of NGS technology has assisted greatly in terms of mapping genetic diversity within *Entamoeba* [13, 33, 116, 182-184], but it is also clear that experience is needed in terms of interpreting such data. Quite a few sequences are quite short and deposited in GenBank as '*Entamoeba* sp.', even in cases where it would be possible to assign both a species name and even a subtype name.

In the following, a selection of 'non-pathogenic' *Entamoeba* species will be accounted for with focus on genetic diversity and host specificity. There will be examples of survey data for each species and these data are included mainly to provide a small impression of the global prevalence of the species.

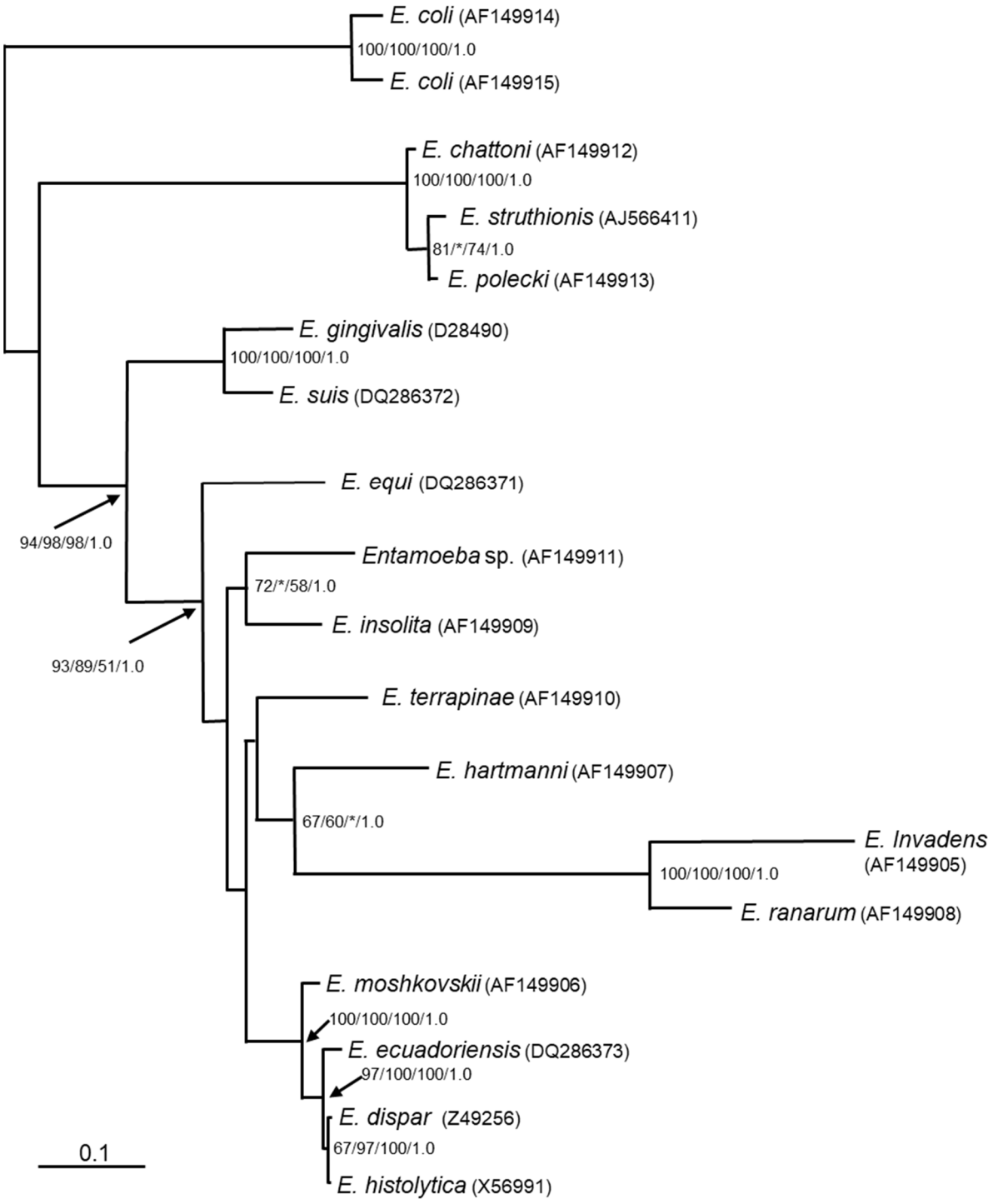

E. coli (AF149914)
100/100/100/1.0
E. coli (AF149915)
E. chattoni (AF149912)
100/100/100/1.0
E. struthionis (AJ566411)
81/*/74/1.0
E. polecki (AF149913)
E. gingivalis (D28490)
100/100/100/1.0
E. suis (DQ286372)
E. equi (DQ286371)
94/98/98/1.0
Entamoeba sp. (AF149911)
72/*/58/1.0
E. insolita (AF149909)
93/89/51/1.0
E. terrapinae (AF149910)
E. hartmanni (AF149907)
67/60/*/1.0
E. Invadens
(AF149905)
100/100/100/1.0
E. ranarum (AF149908)
E. moshkovskii (AF149906)
100/100/100/1.0
E. ecuadoriensis (DQ286373)
97/100/100/1.0
E. dispar (Z49256)
67/97/100/1.0
E. histolytica (X56991)
0.1

Figure 4, A.

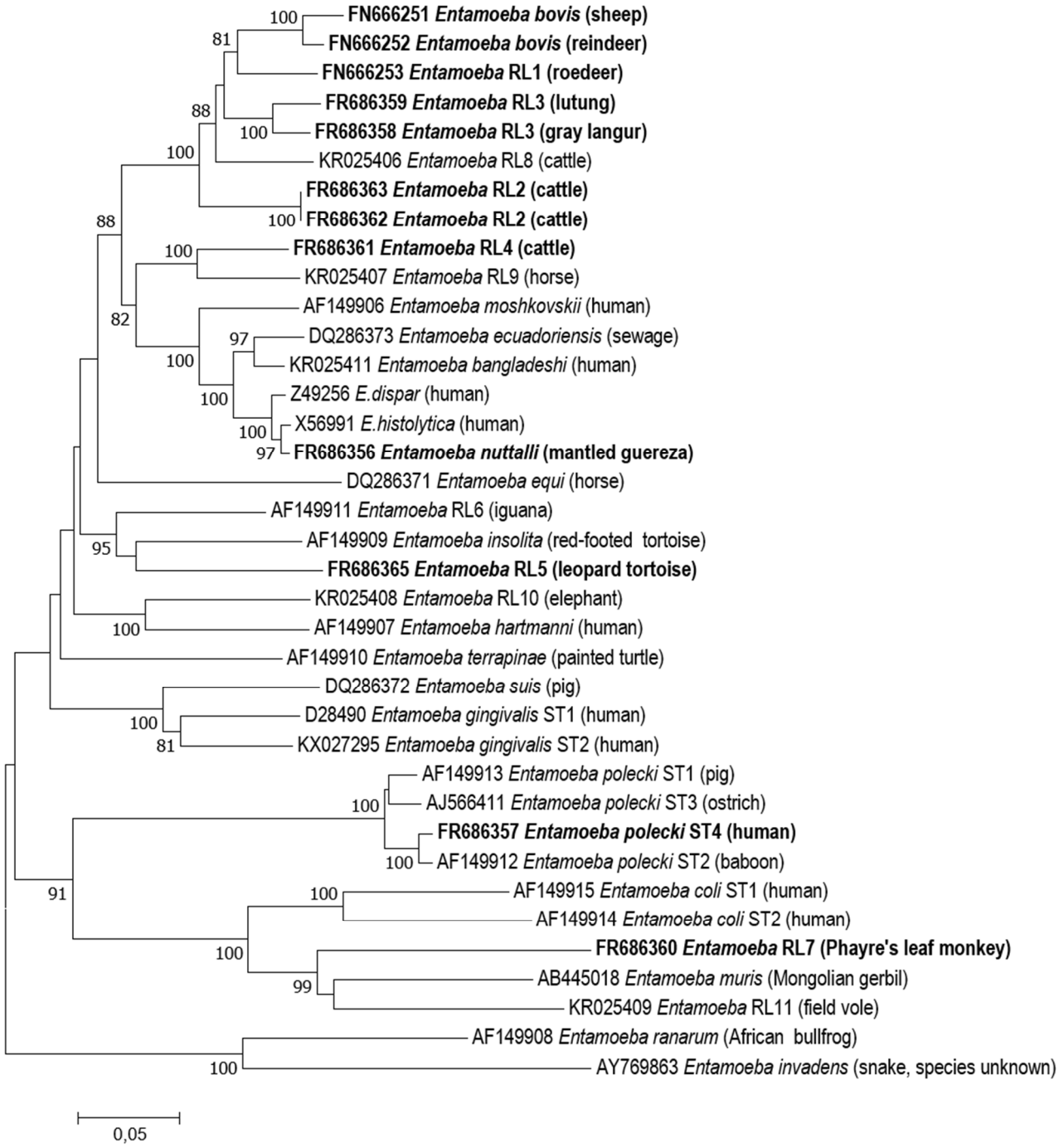


Figure 4, B.

Figure 4, A and B. Example of how *Entamoeba* phylogeny has developed over the past couple of decades. The analysis comprises complete or near-complete SSU rDNA *Entamoeba* sequences. Example A is a reproduction of the tree included in the work by Clark and colleagues from 2006 [185]. In example B, all ribosomal lineages and subtypes reported to date (January 2023) were included (reference sequence depository available here: http://entamoeba.lshtm.ac.uk/ref.entamoeba.txt). The sequences highlighted in bold font are sequences that were published for the first time in the articles shortlisted for this thesis. The tree in part B was generated for the present thesis. It used distance-based analysis (Neighbour-Joining algorithm with 1,000 bootstraps) of 1,526 positions in an alignment of 37 nt sequences. The scale bar indicates nt substitutions per site.

2.3.1. Entamoeba coli

2.3.1.1. Aspects of biology, diagnosis, and epidemiology of *Entamoeba coli*

Octonucleate cysts of *Entamoeba* have been identified in faeces from human and non-human primates as well as from ungulates and rodents. Among these, *E. coli* is a common finding in human faecal samples, with survey positivity rates reaching almost 40% in countries such as Burkina Faso and Venezuela [186-188]. In Denmark, the *E. coli* colonisation rate may reach at least 14.6%, depending on the population studied [90, 105, 189, 190]. If not defeated by *E. gingivalis*, *E. coli* may very well be the most common species of *Entamoeba* to colonise humans, and most of the 26 wastewater samples analysed by metabarcoding were positive [33].

Although cysts of *E. coli* are generally larger than those of *E. histolytica*, Dobell and Jepps [191] claimed that it is impossible to use the size of the cyst alone for the differentiation of these two species. Hence, the number and structure of the nuclei in mature cysts may be the only hallmark to separate these two species morphologically.

Entamoeba coli-specific primers were developed for PCR-based detection and molecular characterisation by Stensvold et al. [49] and later used by for instance Chihi *et al.* to detect and differentiate *E. coli* [192]. A different approach to specific molecular detection of *E. coli* was taken by Matey and colleagues who developed a nested PCR for specific detection [193]. This methodology was used by Matsumura and colleagues in a survey of healthy Indonesian school children, where a positivity rate of 44% was reported [194].

2.3.1.2. Genetic diversity and host specificity of *Entamoeba coli*

A bit more than ten years back, the extensive use of ribosomal gene sequencing from a diverse set of *E. coli*-positive samples enabled us to identify two different lineages within *E. coli*, and this led to the first published evidence of cryptic genetic diversity in *E. coli* with a suggestion to divide the species into subtypes (ST), ST1 and ST2 [49]. Unfortunately, data on cyst morphology were not available for study to identify any differences in cyst size between ST1 and ST2 that could explain the early findings of Dobell [195] and Matthews [196] that suggested a bimodal cyst size distribution among *E. coli* cysts.

When the data were published (2011), it appeared that ST1 was genetically homogenous compared with ST2, and that only sequences from humans were in ST1, whereas ST2 had sequences from both human and non-human primates. Recently, we revisited the genetic diversity of *E. coli* and investigated two different regions of the SSU rRNA gene to minimise the risk of losing important information [44].

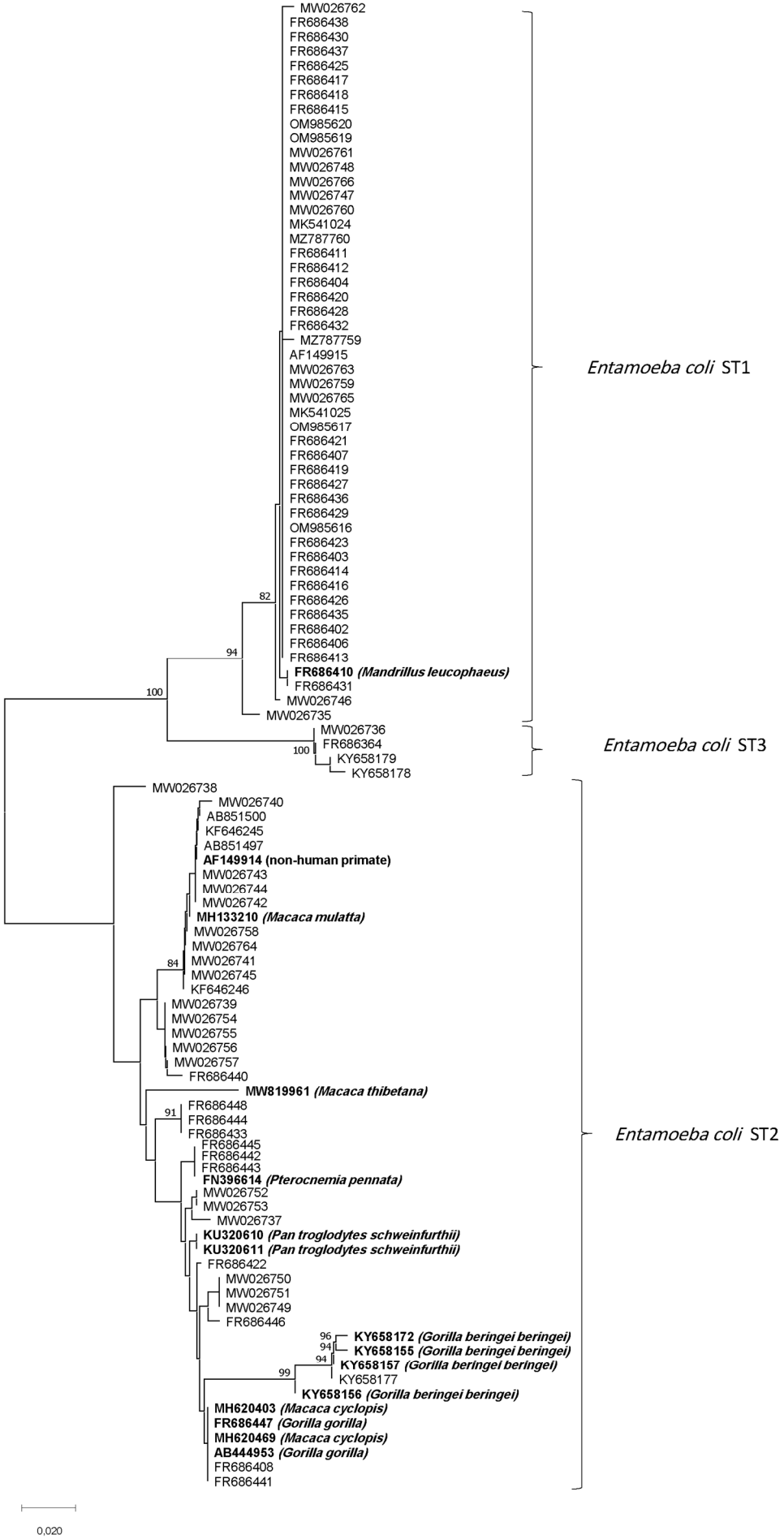

MW026762
FR686438
FR686430
FR686437
FR686425
FR686417
FR686418
FR686415
OM985620
OM985619
MW026761
MW026748
MW026766
MW026747
MW026760
MK541024
MZ787760
FR686411
FR686412
FR686404
FR686420
FR686428
FR686432
MZ787759
AF149915
MW026763
MW026759
MW026765
MK541025
OM985617
FR686421
FR686407
FR686419
FR686427
FR686436
FR686429
OM985616
FR686423
FR686403
FR686414
FR686416
FR686426
FR686435
FR686402
FR686406
FR686413
FR686410 (Mandrillus leucophaeus)
FR686431
MW026746
MW026735
MW026736
FR686364
KY658179
KY658178
MW026738
MW026740
AB851500
KF646245
AB851497
AF149914 (non-human primate)
MW026743
MW026744
MW026742
MH133210 (Macaca mulatta)
MW026758
MW026764
MW026741
MW026745
KF646246
MW026739
MW026754
MW026755
MW026756
MW026757
FR686440
MW819961 (Macaca thibetana)
FR686448
FR686444
FR686433
FR686445
FR686442
FR686443
FN396614 (Pterocnemia pennata)
MW026752
MW026753
MW026737
KU320610 (Pan troglodytes schweinfurthii)
KU320611 (Pan troglodytes schweinfurthii)
FR686422
MW026750
MW026751
MW026749
FR686446
KY658172 (Gorilla beringei beringei)
KY658155 (Gorilla beringei beringei)
KY658157 (Gorilla beringei beringei)
KY658177
KY658156 (Gorilla beringei beringei)
MH620403 (Macaca cyclopis)
FR686447 (Gorilla gorilla)
MH620469 (Macaca cyclopis)
AB444953 (Gorilla gorilla)
FR686408
FR686441
Entamoeba coli ST1
Entamoeba coli ST3
Entamoeba coli ST2
82
94
100
100
84
91
96
94
94
99
0,020

Figure 5 (previous page). Intraspecies variability of *Entamoeba coli* based on 104 DNA sequences representing the 5'-end of the SSU rRNA gene retrieved from the NCBI database. The *E. coli*-specific sequences obtained from non-human hosts are indicated in boldface type. Non-human *E. coli* hosts are indicated in parentheses after the NCBI database ID accession number. Three subtypes are acknowledged (ST1–ST3). The Neighbor-Joining method was used. Evolutionary distances were computed using the Kimura 2-parameter method. All ambiguous positions were removed for each sequence pair (pairwise deletion option). There were 244 positions in the final dataset. Only bootstrap values >70 are shown. The scale bar indicates nt substitutions per site. The figure is reproduced from the work by Stensvold et al., 2022 [44].

Our phylogenetic analyses this time pointed towards not only two but three subtypes (**Figure 5**). This study confirmed the genetic homogeneity of ST1 and disclosed further heterogeneity in ST2. However, a non-human primate *E. coli* sequence appeared in the ST1 clade, namely one from a *Mandrillus leucophaeus* (FR686410), and AB749457 was found in a macaque in China, which means that ST1 is not exclusive to humans. The other NHP sequences clustered with ST2.

To date, a good 200 *E. coli*-specific sequences have been deposited in GenBank, and these are all SSU rDNA sequences (**Table 9**). Across the six near-complete *E. coli* SSU rDNA sequences in GenBank available, up to at least ~12% genetic difference can be appreciated.

When reviewing GenBank data for the development of this thesis, I took note of a 527 bp-long sequence deposited in 2020, namely MW026738 from a human sampled in the Amazonas regions, Brazil. Phylogenetic analysis of this sequence could indicate that it might represent a new subtype of *E. coli*, which in that case would be ST4 (**Figure 6.**). If possible, a longer sequence should be produced for this strain to enable a more robust phylogenetic analysis. But even then, it might not be possible to tell with much confidence whether the sequence would represent a basal branch of ST2 or indeed a new clade. This leaves us in the same position as we were with what is now *E. coli* ST3; originally we had sequence S2702 that appeared as a deep branch at the base of ST1, and that is what we called it [49]. Even though it was a complete sequence it was not enough for us to call it a new subtype until we had additional sequences from multiple samples that clustered together with S2702 [44].

When scrutinizing the relatively large genetic variety within ST2, subclades seem to form within the subtype, and there are examples of specific variants that have been found across the globe in independent studies. One such example is made up by the sequences MW819961, MK559462 and AB749456. These are all from monkeys and share SNPs not seen in other ST2 sequences (data not shown).

In their study of protist parasites in human-habituated mountain gorillas (*Gorilla beringei beringei*), humans and livestock from Bwindi Impenetrable National Park, Uganda, Nolan and colleagues identified a few sequences (KY658155–KY658157, KY658172, and KY658177–KY658179) that they deposited as '*Entamoeba* sp.' in GenBank in 2017 [197]. When included in a phylogenetic analysis, these sequences cluster with *E. coli* ST2 (**Figure 5**) [44].

The ST3 clade holds only a limited number of sequences, which are all from humans, and all of which stem from either South America, Africa, West Bank, or Iraq, including two from human samples from the Uganda study mentioned above [44, 197]. The ST3 data from Iraq have not been included in any publication so far. The vast majority of ST2 sequences identified to date are also from South America or Africa, while ST1 appears to have limited geographical restriction [44]. Studying the genetic diversity of *Entamoeba* in humans in Brazil, Calegar and colleagues found evidence of all three subtypes, including ST3 (MW026736); their data were in support of extensive levels of intra-subtype genetic diversity in *E. coli* [198].

The species name '*E. coli*' has been applied mainly to octonucleated cysts found in faeces from human and non-human primates, and among the 214 sequences included in **Table 9**, 183 *E. coli* sequences are from humans, and 21 from NHPs; the one remaining sequence, for which information on host species is available is from a rodent (see below). However, as early as in 1928, Kessel reported the finding of *E. coli* in a pig [199], and there are recent reports on *E. coli* commonly observed in pigs sampled in Colombia [200, 201]. Unfortunately, no molecular data were included for *E. coli* in the two Colombian studies, and so it remains to be confirmed whether the parasites referred to as *E. coli* were in fact *E. coli*, another octonucleated-cyst producer such as *Entamoeba* RL7, or even other species. *E. coli* cysts with a diameter as small as 11 µm have been reported [191], and those could be difficult to separate from other *Entamoebas* in pigs, such as *E. polecki* or *E. suis*, in case the nuclei cannot be clearly discerned.

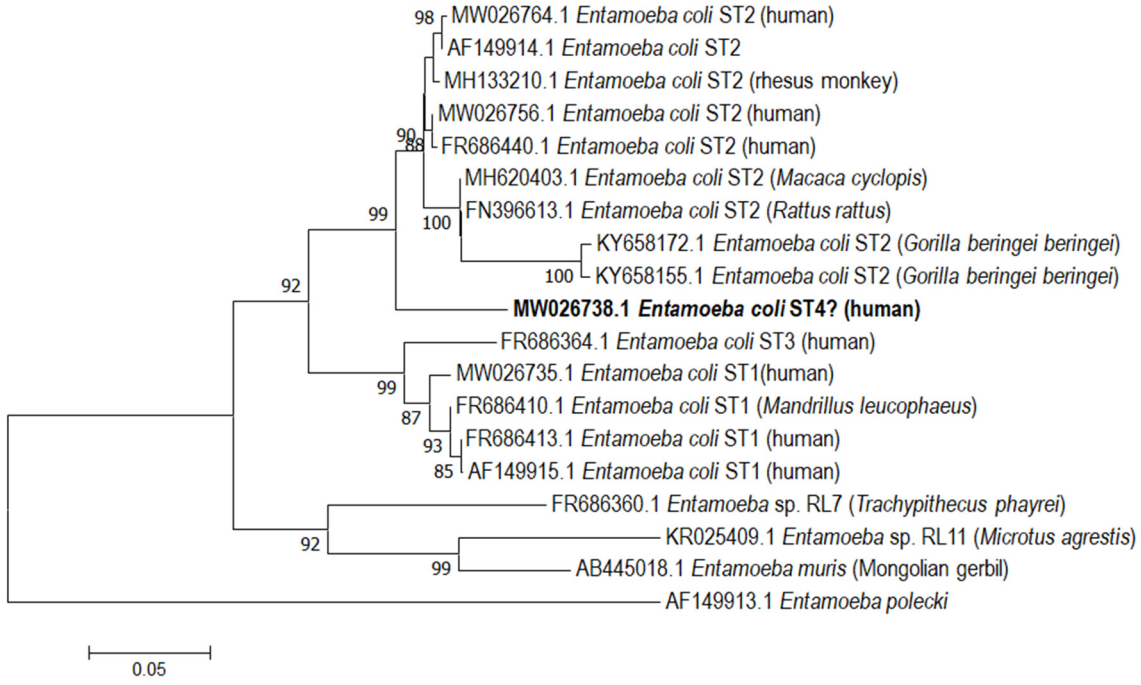


Figure 6. Neighbor-joining analysis (1,000 bootstraps) of the sequence MW026738 (highlighted in boldface) with reference sequences reflecting all three suggested subtypes of *Entamoeba coli* and AF149913 (*Entamoeba polecki*) as outlier. The analysis could suggest that MW026738 represents a novel subtype, which in that case could be referred to as ST4. There were a total of 554 positions in the final dataset. The scale bar indicates nt substitutions per site.

In non-human primates, *E. coli* is a frequent finding, and both apes and monkeys appear to be natural hosts. Positivity rates of 20%–70% have been reported in sun-tailed monkeys [202], macaques [203, 204], baboons [205], gorillas [206, 207], and chimpanzees [208].

Experiments seeking to infect rodents with *E. coli* produced discrepant results. Both Kessel [209] and Regendanz [210] reported successful experimental infection of rodents, whereas Neal [211] failed to be able to establish *E. coli* in mice and rats; meanwhile, experimental infection with *E muris*, which is closely related octonucleate-cyst producer was achieved. It is possible that different subtypes of *E. coli* have been used, and it may be possible that one or more may be able to colonise/infect rodents. Indeed, Ponce-Gordo and colleagues deposited a SSU rDNA sequence in the NCBI database referred to as '*E. muris*' from a rat (FN396613); however, in our phylogenetic analysis, FN396613 clusters with *E. coli* ST2 [44], so a plausible theory would be that ST2 can be hosted by rodents, whereas ST1 cannot. Indeed, we identified *E. coli* ST2 in a chinchilla [49], which could support the hypothesis of ST2 being able to colonise rodents.

In our recent study of Swedish wastewater samples, both ST1 and ST2 were found, with ST1 being more common [33].

Apart from *E. muris*, a couple of other ribosomal lineages clustering with *E. coli*, namely *Entamoeba* RL7 and *Entamoeba* RL11. RL11 was found in a field vole by Jacob et al. [28], while RL7 was identified in both humans and langurs [28, 49]. Given the relatively consistent pattern between phylogenetic topology and number of nuclei in mature cysts, it would appear reasonable to assume that both RL7 and RL11 are octonucleate cyst producers; this has been confirmed for RL7, whereas no morphological data are yet available for RL11.

On the basis of host and cyst morphology, other octonucleate cyst-producing species have been described, namely *Entamoeba cavie*, observed in laboratory guinea pigs; *Entamoeba cuniculi*, a parasite of the rabbit, and *Entamoeba gallinarum* from chicken and turkeys [212].

The host specificity, distribution and genetic diversity of *E. coli* subtypes call for further investigations to delineate the role of this species in human health and disease and to identify routes of transmission.

2.3.2. Entamoeba dispar

2.3.2.1. Aspects of biology, diagnosis, and epidemiology of *Entamoeba dispar*

The species *E. dispar* was introduced by Emilie Brumpt in 1925 after recognising that an organism that could not morphologically be differentiated from *E. histolytica* did not give rise to intestinal symptoms in colonised individuals [181]; moreover, it did not produce invasive disease in cats, which at that time were typically used in studies to prove the invasiveness of *E. histolytica*. His observations were largely ignored over the next many decades, possibly due exactly to the fact that these two species could not be told apart.

It was only about 50 years later that isoenzyme analysis and molecular methods assisted in providing evidence in support of Brumpt's observations [20, 21, 213-216]. Sequencing of the near-complete SSU rRNA genes of *E. histolytica* and *E. dispar* revealed a genetic difference between the two of ~1.5%. After several studies focussing on the distribution of and ways to separate the two species, Verweij and colleagues developed a real-time PCR for simultaneous detection of *E. histolytica*, *Giardia lamblia*, and *Cryptosporidium parvum* in faecal samples using multiplex real-time PCR in 2004 [217], which heralded a new area in the diagnostic parasitology laboratory. This assay took advantage of the fact that *E. histolytica* and *E. dispar* could be separated based on SSU rRNA genes, and so the 22 bp-long *E. histolytica*-specific TaqMan probe was developed so that it annealed to a relatively polymorphic region (no less than six mismatches) separating *E. histolytica* from *E. dispar*. This type of strategy and methodology informed the development of numerous molecular analyses now in place in most modern parasitology laboratories [11].

E. dispar appears to be quite a common parasite of cosmopolitan distribution. In a survey of 199 healthy school children sampled in Nigeria, we identified a positivity rate of 18.6% [218]. Among human immunodeficiency virus (HIV)-infected patients followed up in Denmark, the positivity rate using the same method (real-time PCR) was 10.4% [105]; however, in a random subset of samples tested in our laboratory, we only saw a positivity rate of 2/889 (0.2%) [219], indicating that HIV-infected individuals might be a population particularly prone to developing *E. dispar* colonisation compared with non-HIV infected individuals, at least in this country. To this end, we identified *E. dispar* in only one of 41 Syrian asylum seekers in Denmark [189]. We found *E. dispar*-specific DNA in 11/26 (42.3%) wastewater samples from Sweden [33].

Among 175 human patients with intestinal symptoms sampled in Egypt, the positive rate was 40.5% by PCR [220], and data from a study from South Africa indicated the presence of *E. dispar* in 14.7% of 170 patients tested [221]; however, in both studies, the diagnosis of *E. dispar* was done only by gel inspection of PCR products with no confirmatory DNA sequencing.

There is one DNA sequence in the NCBI database (KX357142; **Table 10**) deposited by a team in India that is from a pus sample from a human. In the metadata, the title 'Amoebic Liver Abscess caused by *Entamoeba dispar* and *Staphylococcus aureus*' has been provided, but to my knowledge, no such article has been

published to date. There are additional sporadic sources of information that could indicate that some strains of *E. dispar* could be pathogenic [222-224].

2.3.2.2. Genetic diversity and host specificity of *Entamoeba dispar*

To my knowledge, the species has been identified mainly, if not only, in primates. In non-human primates, positivity rates from surveys have reached at least 18.1% in mixed study populations of apes and monkeys [225]. Pomajbikova and colleagues found a positive rate of 16% specifically in *Pan troglodytes schweinfurthii* [208]. There are a limited number of sequences (OP453103–OP453107) in the NCBI database for which 'human and dogs' have been entered in the 'host' field, and so it remains unknown to most whether these sequences are from humans and/or from dogs in the absence of an accompanying article. Data that could indicate an even much higher prevalence among NHPs are those published by Dong and colleagues in their impressive study from China [226]; however, the exact positivity rates are difficult to decipher in that article. It could be speculated that NHPs may constitute a reservoir for human carriage of *E. dispar*.

To date, at least 134 SSU rDNA sequences of *E. dispar* are available in GenBank (**Table 10**). Although the within-species diversity of *E. dispar* appears to be limited (~0.5%) and confined to a handful of SNPs, these are organised in a way that could fuel a hypotheses of the existence of at least two separate subtypes (**Figure 7.**). There is currently, however, limited data, so it may be premature to hypothesise on differences in geographic distribution and cryptic host specificity.

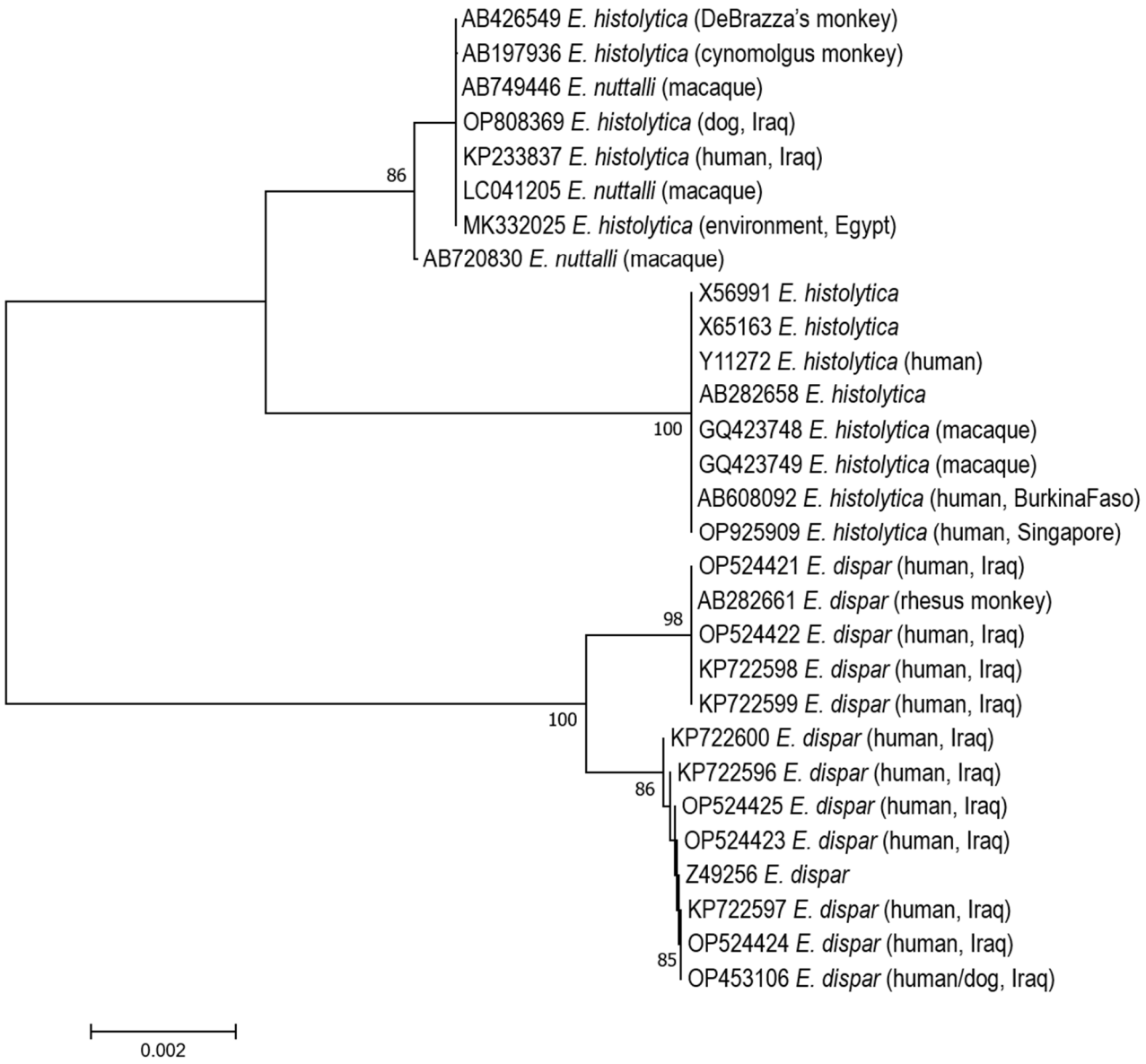


Figure 7. Phylogenetic analysis of SSU rDNA sequences of *Entamoeba histolytica*, *Entamoeba dispar*, and *Entamoeba nuttalli.* The tree was generated by Neighbour-Joining method-based analysis of 1,658 unambiguously aligned nt positions, corresponding to about ~90% of the SSU rRNA gene. Information on host and geography is provided where available in GenBank. Three clades can be appreciated: The upper clade is separated from the remaining sequences by a bootstrap support of 86 and holds *E. nuttalli* sequences (although some were deposited in GenBank as '*E. histolytica*'). The second clade holds *E. histolytica* (*sensu stricto*) sequences and is supported by a bootstrap value of 100. The third clade is also supported by a bootstrap value of 100 and holds sequences specific to *E. dispar*. The *E. dispar* clade again holds two clusters, which could indicate the presence of two subtypes of *E. dispar.* The genetic distance between the two clusters is up to about 0.5%. The scale bar indicates nt substitutions per site.

2.3.3. Entamoeba gingivalis

2.3.3.1. Aspects of biology, diagnosis, and epidemiology of *Entamoeba gingivalis*

Being a digestive-tract *Entamoeba* species with no evidence of a cyst stage, *E. gingivalis* is a parasite of the oral cavity passed on by direct buccal contact and/or by saliva or contaminated mouth utensils or food [227, 228]. As the name implies, this amoeba lives in the gingival areas around the teeth. The textbooks of Levine [212] and Noble and Noble [229] concur on *E. gingivalis* being a harmless commensal, although "often present in diseased gums".

The parasite appears to be a quite common finding and may be the most common *Entamoeba* species colonising humans. PCR-based testing of saliva and dental plaques revealed colonisation in 1 of every 5 women sampled in Iran [230]. Also in Iran, PCR testing revealed an overall positivity of rate of 11.7% in randomly selected adolescents, with colonisation being statistically significantly linked to i) a gingival index that indicated severe inflammation and ii) having decayed, missing, and filled teeth [228]. In neighbouring Turkey, and also using PCR, Yaseen and colleagues identified colonisation rates of 88.9%, 84.9% and 47.9% in patients with periodontitis, patients with gingivitis and in healthy individuals, respectively [231]. Until recently, there was a lack of data from the 'background population'. However, in our recent survey of Tanzanians with non-oral/non-dental diseases (N = 52), 31% of the study individuals tested positive for *E. gingivalis*-specific DNA extracted from oral washings [43].

Recently, Keeler and colleagues identified *E. gingivalis* as a likely host of human-associated redondoviruses, which have a high prevalence in healthy humans, but the abundance of which is increased in patients with periodontitis, acute illness and severe Coronavirus Disease 2019 [232]. We and others have investigated associations between bacterial communities and the presence/absence of *E. gingivalis* in mouthwash samples, subgingival plaque or other types of samples from the oral cavity. In our laboratory, we found that, despite higher microbial diversity in *E. gingivalis* carriers, the top-ten most common bacterial genera were almost similar; only *E. gingivalis* carriers were more likely to be colonised by *Aggregatibacter* [43], which has been associated with periodontal disease. Moreover, *Neisseria* spp. were enriched in carriers relative to non-carriers. These observations confirmed those of Koller and colleagues [233], who speculated that *E. gingivalis* might promote oral cavity colonisation by phagocytosis-resistant bacteria [233].

2.3.3.2. Genetic diversity and host specificity of *Entamoeba gingivalis*

At least two subtypes, ST1 and ST2, have been observed. In 2018, García and colleagues introduced evidence of a 'new subtype' of *E. gingivalis*, which was named '*E. gingivalis* ST2, kamaktli variant' (*e.g.,* KX027297) [234], and which only shared 89% similarity with another *E. gingivalis* sequence (KX027298); the latter sequence was referred to a ST1 and had high similarity (99.58%) to D28490, originally deposited back in

1995 by Yamamoto and colleagues [235]. This new variant was identified in patients with dento-oral diseases in Mexico. The D28490 sequence had been obtained from an ATCC (American Type Culture Collection) strain from a subgingival space of adult with periodontal disease (geographical data lacking). The authors suggested that this new variant could in fact be the SSU rDNA sequence of one or more of the species *Entamoeba pyogenes*, *Entamoeba canibuccalis*, *Entamoeba equibucalis*, and *Entamoeba suisginvalis*, but this is not possible to confirm, since these species names were introduced in the absence of sequence data.

Interestingly, in August 2022, rDNA (18S-ITS1) sequences of *E. gingivalis* were published from a study of individuals sampled in Austria. Here, the research team found evidence of what in GenBank is proposed as a new subtype, ST3. However, the publication linked to the GenBank accession numbers (**Table 11**) is not yet available. Nevertheless, I include a phylogenetic analysis that is indeed in support of the new subtype (**Figure 8**). Since the sequences from Austria only included the very 3'-end part of the SSU rRNA gene along with the ITS1 region (when García and colleagues launched the idea of the existence of two subtypes, they based it on complete SSU rDNA + ITS analysis), it is currently unknown, how much ST3 differs genetically from the two other subtypes across the SSU rRNA gene alone.

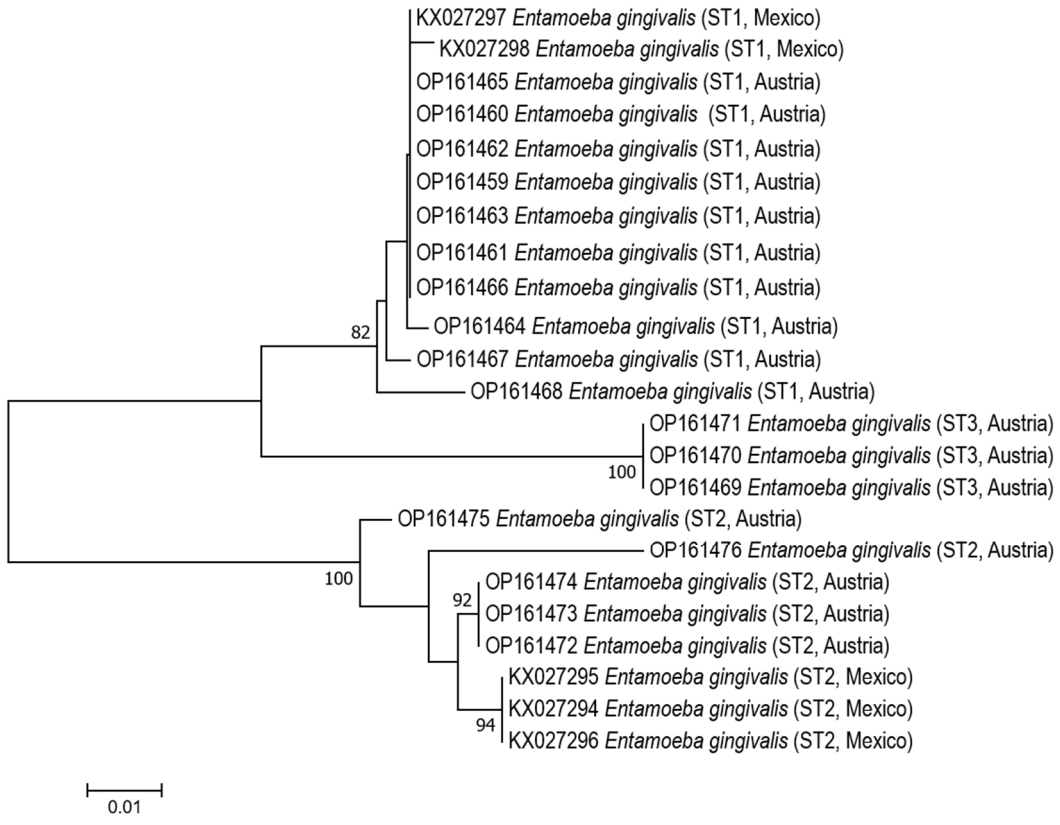


Figure 8. Phylogenetic analysis of rDNA sequences of *Entamoeba gingivalis* reflecting the 3'-end of the 18S + the ITS1 region. Three subtypes can be appreciated. The tree was generated by neighbour-joining analysis of 348 unambiguously aligned nt positions. The evolutionary history was inferred using the Neighbor-Joining method and the evolutionary distances were computed using the Kimura 2-parameter method. The scale bar indicates nt substitutions per site.

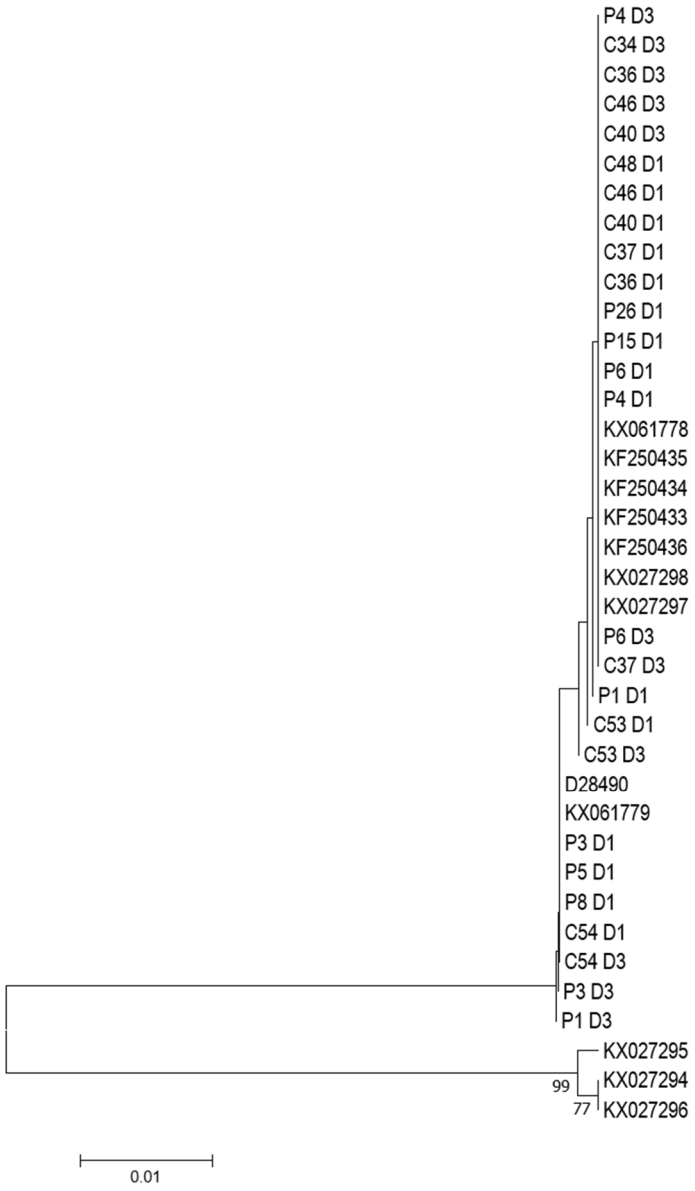


Figure 9. Phylogenetic analysis of SSU rDNA sequences of *Entamoeba gingivalis.* The tree was generated by neighbour-joining analysis of 330 unambiguously aligned nt positions, corresponding to about ~20% of the SSU rRNA gene. Sequences without GenBank accession numbers are sequences obtained in our study from Tanzania [43].Two clades can be appreciated; the top clade with the majority of the sequences represents ST1, and the lower clade (with three sequences, which are all reference sequences) represents ST2. The scale bar indicates nt substitutions per site.

The intra-subtype diversity within the three subtypes identified today remains relatively unexplored given the few studies that involve molecular characterisation of *E. gingivalis* (probably only a dozen of articles according to PubMed). In our study of *E. gingivalis* sequences from the Tanzanian mouthwash samples, which all belonged to ST1, a single SNP was identified, and this SNP was also picked up by García and colleagues [234]. An overview of the sequences obtained in our Tanzanian study is provided in **Figure 9**. As seen, some

of the consensus sequences for test individuals for whom two consecutive sequences were available (reflecting two different time points of sampling), were not completely identical (*e.g.,* the sequences for the individuals with the IDs 'P1' and 'C53'). However, it should be borne in mind here that the consensus sequences were generated from PCR products sequenced by ILLUMINA sequencing technology and that for some samples, the number of sequence reads available might have been quite small. This might have given rise to one or two SNPs across the consensus sequence that would be due to sequencing error rather than 'true' polymorphism.

In terms of relative distribution, data are still scarce, but ST1 appears to be the predominant version of *E. gingivalis* [236]. Nevertheless, ST2 has been found in both Mexico [234], Turkey (Orsten *et al.,* unpublished GenBank entries OP456215 and OP422447), and in DNA from bronchioalveolar lavage samples in Denmark (Stensvold *et al.,* unpublished observations) suggesting a potentially global distribution.

The level of genetic diversity within *E. gingivalis* is reminiscent of the level observed within *E. coli* (between 10% and 15%), and it is highly likely that it would be useful to consider the three subtypes as three distinct species.

A PubMed search on *E. gingivalis* in dogs renders two articles [237, 238], while none when cats are searched as hosts. For now, the host spectrum of *E. gingivalis* appears to be extremely narrow. To date, 50 SSU rDNA sequences have been deposited in GenBank (**Table 11**), all being from humans.

Future research might reveal whether the two subtypes differ in terms of clinical significance and epidemiological features.

2.3.4. Entamoeba hartmanni

2.3.4.1. Aspects of biology, diagnosis, and epidemiology of *Entamoeba hartmanni*

Entamoeba hartmanni is a quadrinucleate cyst-producing amoeba, with cysts resembling those of *e.g., E. histolytica*, apart from the fact that they tend be smaller, with a size of approximately 8 µm in diameter [212], and with the peripheral chromatin of the nuclei tending towards being arranged in lumps [212]. For many years, this species was confused with *E. histolytica* (considered 'small-race *E. histolytica* [239-241]), and so, data from surveys based on morphology only should be interpreted with caution. Generally, however, *E. hartmanni* appears to be a less common than *E. coli*; not only in humans but also in non-human primates.

Matsumura used nested PCR followed by species-specific amplification of species of *Entamoeba*, identifying a positivity rate of 31% in healthy school children in Indonesia [194] and an association between *E. hartmanni* colonisation and loose stools; it should be mentioned, that no sequence data were available for this study, so it remains unknown which subtype(s) (see below) might have been involved. *Entamoeba hartmanni* generally appears to be less common than *E. coli*. In Denmark, positivity rates in different study populations based on

traditional microscopy have been found to range between 0.8% and 7.3% [90, 189], and in our metabarcoding analysis of Swedish wastewater samples, *E. hartmanni* was one of the least common parasites detected, with a positivity rate of 15% [33].

2.3.4.2. Genetic diversity and host specificity of *Entamoeba hartmanni*

When we did our first study on molecular characterisation of *E. hartmanni* from humans and NHPs [49], we did not obtain data that could indicate the existence of more than one species; neither did other authors after us [13, 33, 198]. In our most recent study, however, which involved characterisation of a part of the gene that was different from the one previously investigated, we found evidence pointing towards the existence of three subtypes (**Figure 10**) [44].

At the time of writing, there are at least 84 SSU rDNA sequences of *E. hartmanni* in the NCBI database, 64 of which are from humans (**Table 12**). There are 10 and five sequences from NHPs and domestic dogs, respectively, and for five sequences, no information on host is available. The five sequences from dogs were published only in October 2022, but no accompanying article is available to date. The data stem from Iraq (Accession numbers OP688358-OP688382), and although Sanger sequencing is indicated as the sequencing method, we do currently not know which type of DNA amplification (including any pre-DNA extraction procedures) that was used to produce the data.

DNA-based evidence of *E. hartmanni* was moreover found in pigs by our team in a recent study of intestinal parasites shared between humans and pigs [13], and indeed, the *E. hartmanni* sequences obtained from pig samples were identical to *E. hartmanni* sequences produced from human stool samples. More specifically, all 10 pigs that were found positive had *E. hartmanni* similar to FR686375, a ST1 sequence from a human. This was the first report of a finding of *E. hartmanni* in pig faeces, thus expanding the host spectrum for this species. In 1928, Kessel [199] reported *Entamoeba* cysts with 1–4 nuclei in pigs, and the cysts were between 5 and 12 μm, which would include the *E. hartmanni* cyst size range.

Studies of NHPs have identified both apes and monkeys as what would appear to be natural hosts. Positivity rates of 11%, 27%, and 51% have been found in *Macaca cyclopis* in Taiwan [242], in *Gorilla gorilla beringei* (mountain gorillas) sampled in Rwanda [207], and in *Pan troglodytes schweinfurthii* in Tanzania [208], respectively; molecular methods were included in the studies from Taiwan and Tanzania.

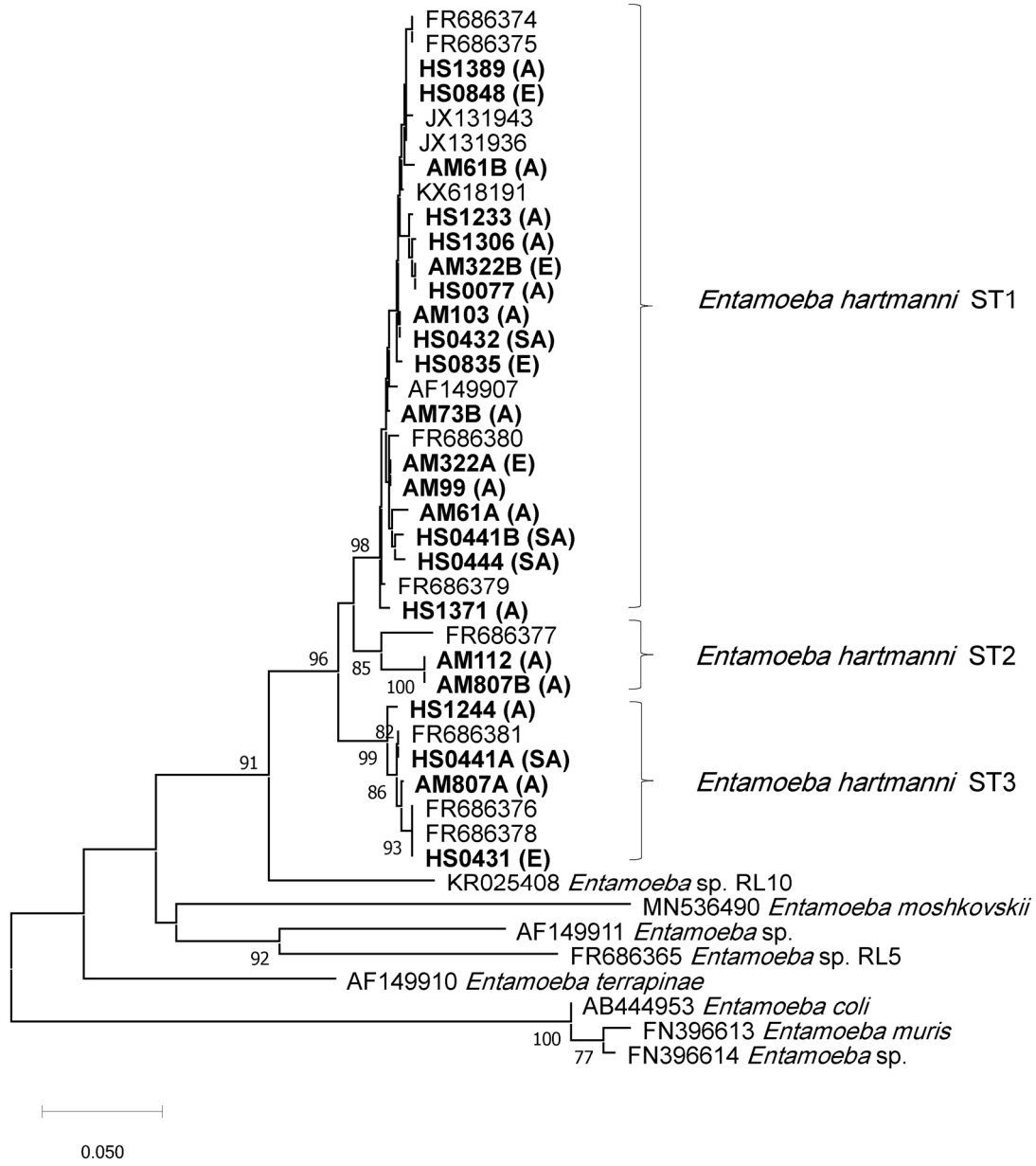


Figure 10. Phylogenetic analysis of *Entamoeba hartmanni*-specific sequences generated in the study by Stensvold *et al.* [44] revealing the existence of three subtypes (ST1–ST3). The Neighbor-Joining method was used. Evolutionary distances were computed using the Kimura 2-parameter method. The analysis involved 43 nt sequences. There was a total of 411 positions in the final dataset. The region covered corresponded to the middle part of the SSU rRNA gene (as opposed to the sequences included in Figure 11, which reflected the 5′-end of the gene, and which does not enable differentiation of subtypes). Only bootstrap values >70 are shown. Sequences generated in the present study are highlighted in boldface and were all from humans. The scale bar indicates nt substitutions per site. Abbreviations used in parentheses for sample origin are as follows: A, Africa; E, Europe; SA, South America.

In the previously mentioned study on protist parasites in human-habituated mountain gorillas in Uganda, Nolan and colleagues [197] deposited a number of ~500 bp-long sequences in GenBank that cluster together with *E. hartmanni* (**Figure 11**) [44]; however, as these are currently referred to as '*Entamoeba* sp.' in GenBank, these sequences were not included in **Table 12**.

Hence, based on DNA data, three major groups of hosts of *E. hartmanni* have now been identified, namely primates, pigs, and dogs. The last group may be a bit surprising, since *Entamoeba* colonising the distal part of the digestive tract is rarely seen in canids. Further sampling of canines and other carnivores is warranted for confirmation.

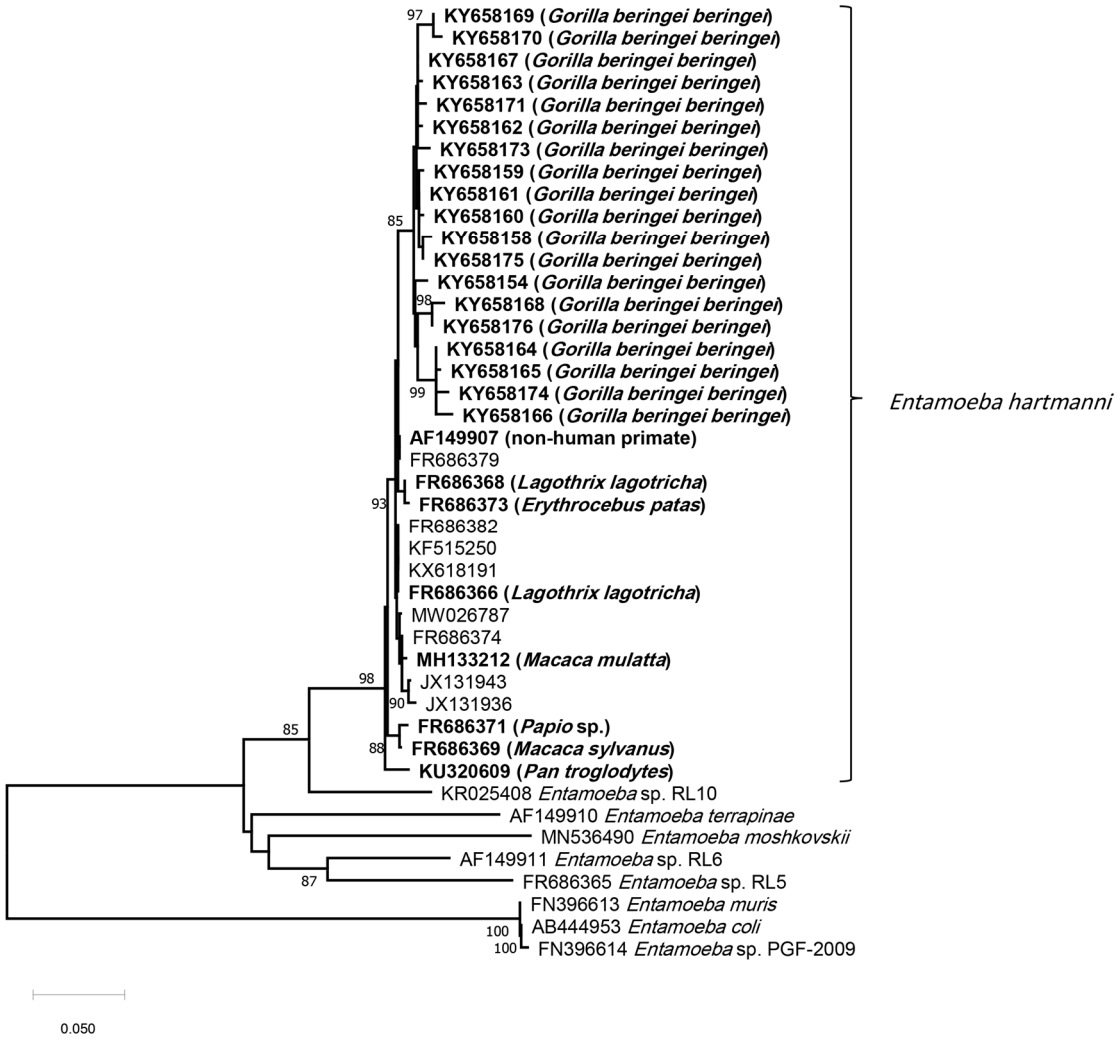


Figure 11. A number of ~500 bp-long SSU rDNA sequences reflecting Entamoebas found in mountain gorillas in Uganda were deposited in GenBank by Nolan and colleagues in 2017 [197]. These cluster with *E. hartmanni*, but since the sequences reflect the 5'-end of the SSU rRNA gene, it was not possible to identify which *E. hartmanni* subtype they belong to [44]. The same applies to the sequences JX131936 and JX131943 found in a study by Hamad and colleagues [243]. The scale bar indicates nt substitutions per site. The tree was reproduced from the study by Stensvold et al., 2022 [44].

2.3.5. Entamoeba moshkovskii

2.3.5.1. Aspects of biology, diagnosis, and epidemiology of *Entamoeba moshkovskii*

Since its description in 1941, *Entamoeba moshkovskii* has been identified mostly in the environment. Reports on a potential link between *E. moshkovskii* colonisation and the development of GI symptoms such as dysentery exist [244, 245]; however, since the species has been known for more than 80 years, and since its pathogenicity still remains to be demonstrated, I have included the species in this account of CLIPPs.

We have not identified *E. moshkovskii* in any human sample by any of the methods in use in our lab for the almost 20 years that I have been working in the Laboratory of Parasitology at SSI; nor do I remember finding it in samples from any non-human host (NHPs, pigs, cattle, muskox, deer, and rodents). Contrasting my own personal experience are the many reports on human *E. moshkovskii* colonisation/infection from 'warmer' parts of the world. López and colleagues reported a positivity rate of 25.4% among children sampled in a district in Cundinamarca, Colombia [246], and even higher rates have been identified in Australia [245]. The ubiquity of the species in sewage samples from Scandinavia coupled with absence of colonisation of mammals in this region is intriguing when compared with the situation in other climate regions, where mammalian colonisation appears to be common, such as India and Bangladesh [244, 247-249], and where there are examples of studies not identifying *Entamoeba* in sewage samples [250]. What are the factors driving mammalian colonisation? Also, while most *Entamoeba* species seen in humans have also been documented extensively in NHPs, to my knowledge there is no DNA sequences from *E. moshkovskii* found in NHPs available in GenBank.

Since the morphology of *E. moshskovskii* is overlapping with that of other quadrinucleate Entamoebas infecting and colonising humans, quite a few DNA-based methods have been developed to facilitate differentiation between these [16, 251-253]. We and others have found NGS-based methods useful to detect and differentiate *E. moshkovskii* from other Entamoebas present in complex matrices [16, 33, 184].

2.3.5.2. Genetic diversity and host specificity of *Entamoeba moshkovskii*

In the book 'Protozoan Parasites of Domestic Animals and of Man', Levine wrote as follows, "This species occurs in sewage. It is not a parasite of animals, but of the municipal digestive tract" [212], indicating that this species would be free living only. Indeed, in our study of Swedish wastewater samples, we identified in *E. moshkovskii* in all of the 26 sample analysed [33], which not only documents the presence of *E. moshkovskii* even in Northern Europe but also points to its ubiquity in "the municipal digestive tract". Nevertheless humans, cattle and turtles have been identified as hosts [28], and there are *E. moshkovskii* SSU rDNA sequences in the NCBI database from pigs (*e.g.,* MW926950), a snake (MN536488), green June beetle (MN536495) and from American cockroaches (*e.g.,* MN536492) (**Figure 12**). A number of near-complete

SSU rDNA sequences have been deposited in GenBank from human stool in Iraq, which sequences are allegedly similar to sequences found in faeces from dogs in Iraq; these sequences are included in **Figure 12**.

It therefore seems likely that this *Entamoeba* is both free living and parasitic. Studies demonstrating *E. moshkovskii* in culture from humans and other hosts would be helpful for confirmation.

At the time of writing, 267 SSU rDNA sequences of *E. moshkovskii* are available in NCBI's nucleotide database (Table 13). Only 26 of these represent sequences longer than 1.45 kb, thereby covering ~80% of the gene or more (**Figure 12**), and the vast majority of these are from either USA or Iraq.

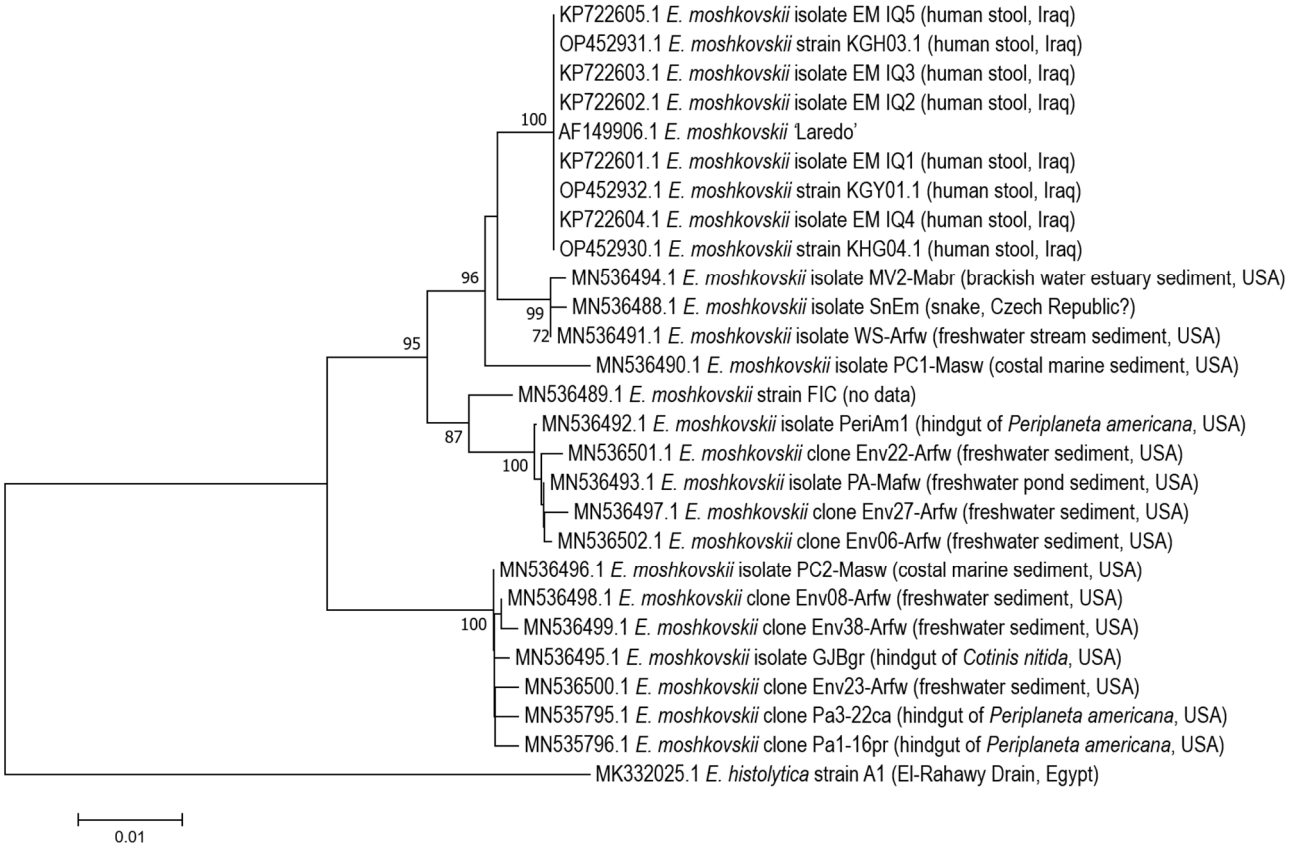


Figure 12. Genetic diversity of *Entamoeba moshkovskii* based on sequences available in the NCBI nt database covering ~80% of the SSU rRNA gene or more (n = 26). Two major clades can be appreciated. There were a total of 1,362 positions in the final dataset. The evolutionary history was inferred using the Neighbor-Joining method (Kimura 2-parameter method). The scale bar indicates nt substitutions per site.

In our study of wastewater samples, we identified two genetic variants of *E. moshkovskii* differing by 6.7%, which indicates substantial genetic diversity within the species [33], and one of which was novel. One variant (MN498050) was found across all the samples tested, and the consensus sequence shared 99.83% similarity with sequences such as MN536502 from freshwater sediment from USA and MN536492 from a cockroach sampled in the USA. The novel variant (MN498051) shared only 94.07% similarity with strains such as KP722601 isolated from human stool in Iraq.

Of course, one of the important questions here is whether environmental strains are genetically different from the strains colonising humans, or at least whether human (and other mammalian) strains are in the same clade(s). Based on our analyses, two major clades exist, and both clades contain strains of both environmental and animal (including human) host origin; however, sequences from humans belonging to the lower clade in **Figure 12** are yet to be identified, and sequences of human and porcine origin cluster primarily with the 'Laredo' strain (data not shown). These observations could easily indicate cryptic host specificity in *E. moshskovskii*.

2.3.6. Entamoeba polecki

2.3.6.1. Aspects of biology, diagnosis, and epidemiology of *Entamoeba polecki*

Prowazek was the first to describe *Entamoeba polecki*, and two years later, in 1914, Swellengrebel introduced the species *Entamoeba chattoni*, which he had observed in NHPs [254].

Entamoeba polecki was considered synonymous with *E. suis* by Levine among others [212]. Meanwhile, in 2006, Clark and colleagues provided evidence of a uni-nucleate cyst producer observed in pig faeces that was genetically much more related to *E. gingivalis* than to *E. polecki*, indicating that suids may host at least two species of uni-nucleate cyst producers namely *E. suis* (Hartmann, 1913) and *E. polecki* [185].

This species is much less commonly reported than other species of *Entamoeba* of similar size. Whether this is due to factual differences in distribution or due to morphological confusion remains unclear. Advanced morphological experience is required to tell the cysts of *E. polecki* apart from those of Entamoebas of similar size, and DNA-based detection and differentiation appear more and more relevant, as more and more laboratories turn to molecular diagnostics. It is doubtful, however, that even modern laboratories screen for *E. polecki* specifically, but these may be picked up by metabarcoding methods, for instance, although we failed to detect *E. polecki* by metabarcoding in three wastewater samples that were identified positive by *E. polecki*-specific PCR and sequencing.

Surveys involving the reporting of *E. polecki* in humans are very scarce. Based on microscopy of faecal concentrates, Desowitz and Barnish identified a positivity rate of 19% among 184 children sampled in Papua New Guinea [255], whereas Park and colleagues reported a positive rate of 1.1% among children sampled in Bat Dambang in Cambodia [256].

2.3.6.2. Genetic diversity and host specificity of *Entamoeba polecki*

Nearly 100 years after the description of *E. polecki*, Verweij and colleagues proposed a new terminology for *E. polecki* [257], which involved the recognition of four subtypes (ST1–ST4) within the species (**figures 13 and 13)**; this initiative appeared relevant due to the overlap in morphology and host specificity (see **Table 14**).

All of the four subtypes have been observed in humans. *E. polecki* ST2 was previously known as *E. chattoni*, and *E. polecki* ST3 has been referred to as *E. struthionis* [49, 53, 257]. The four subtypes differ genetically by up to at least 5% across the SSU rRNA gene.

In 2018, we studied SSU rDNA sequences from 18 stool samples that had been diagnosed with uninucleate *Entamoeba* cysts at the Public Health Agency of Sweden [53]. These sequences were obtained by applying a single-round PCR to genomic DNA extracted directly from stool, using a primer pair (UNINUC_400F and UNINUC_1050R) designed to detect and differentiate subtypes of *E. polecki* by sequencing of the PCR products. Using the same PCR method, we also obtained sequences from six pig faecal samples and three environmental samples (wastewater). Mainly ST4 was seen in the samples from humans, but ST2 and ST3 were seen in a couple of instances. The six pig samples mostly had mixes of ST1 and ST3. Notably, ST4 was not found in any of the three wastewater samples; here, ST1–ST3 were found, with one sample being positive for two subtypes (ST1, ST3).

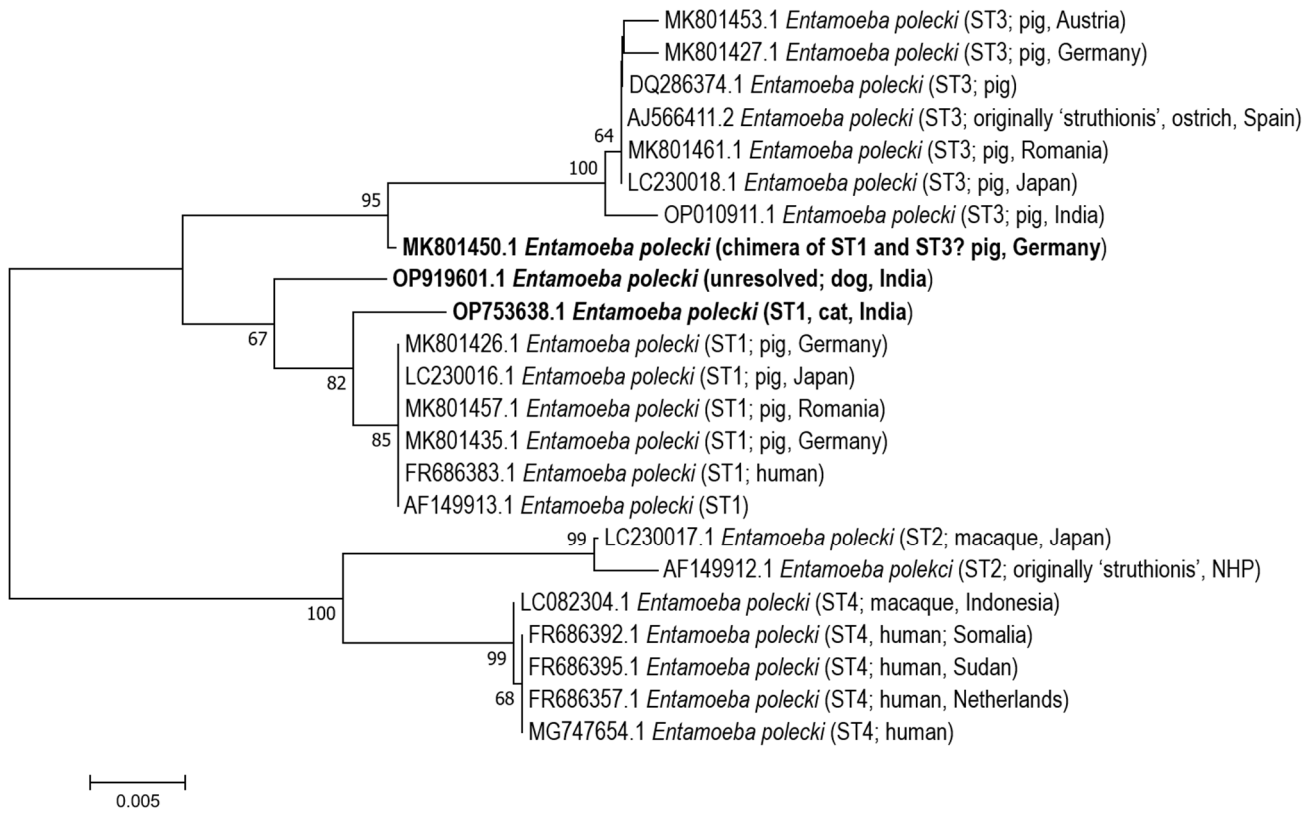


Figure 13. Phylogenetic analysis of *Entameoba polecki* sequences recently published in GenBank (boldface type) together with a selection of *E. polecki* reference sequences representing the four subtypes acknowledged to date (1,000 bootstraps). The Neighbor-Joining method was used. MK801450 is likely a chimaera (see text for details), which could explain why it appears as a deep branch of ST3. In Figure 14, the analysis was repeated without MK801450 to try to obtain better resolution of OP9019601 (sequence from dog) and OP753638 (sequence from cat).There were a total of 562 positions in the final dataset. The scale bar indicates nt substitutions per site. Only bootstrap values >60 are shown. Information on host and geographical origin is provided where known.

The host range of the four subtypes of *E. polecki* is summarized in **Table 14** and overview of *E. polecki*-specific SSU rDNA sequences currently available in the NCBI database is available in **Table 15**. Briefly, ST1 and ST3 have been observed mainly in humans and pigs; ST2 mainly in humans and NHP; and ST4 mainly in humans.

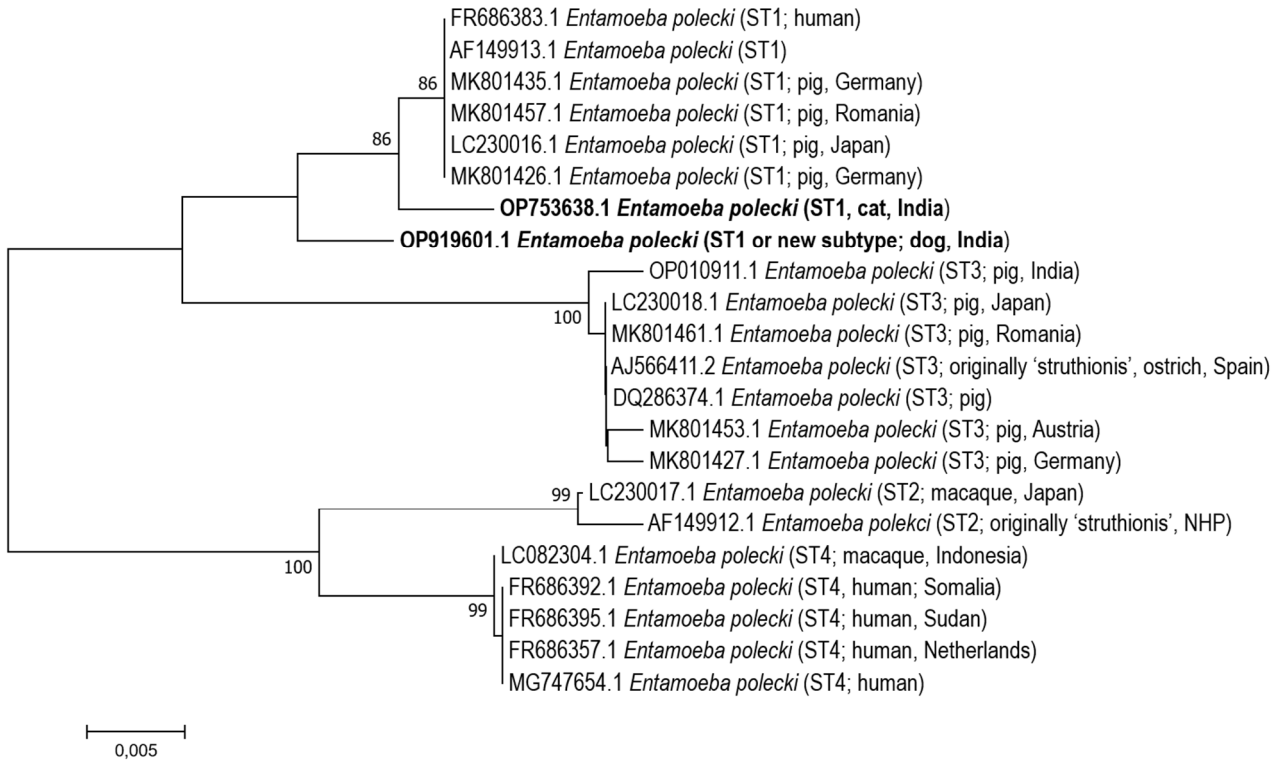


Figure 14. Neighbor-joining analysis of two *Entameoba polecki* sequences recently published in GenBank (boldface type) together with a selection of *E. polecki* reference sequences representing the four subtypes acknowledged to date (1,000 bootstraps). The phylogenetic analysis is similar to that in Figure 13, apart from the fact that MK801450 was not included, since it was likely a chimaera (see text for details). The present analysis was repeated without MK801450 to try to obtain better resolution of OP9019601 (sequence from dog) and OP753638 (sequence from cat). There were a total of 562 positions in the final dataset. The scale bar indicates nt substitutions per site. Only bootstrap values >70 are shown. Information on host and geographical origin is provided where known.

Interestingly, two "*Entamoeba polecki*" sequences were recently deposited as OP919601 and OP753638, originating from a dog and a cat, respectively, sampled in India (**Table 15**). Moreover, a sequence from a pig sampled in Germany was deposited in 2019 as MK801450 (**Table 15**). The latter sequence is likely a

chimaera (a sequence artefact[2]), as it would appear to match AJ566411 (ST3) at both ends but AF149913 (ST1) in the middle, and the hypothesis of it being a chimaera is supported by the fact that it sits on a long branch in **Figure 13**, looking like a deep branch of ST3 that could potentially distort the position of the OP sequences. The MK801450 sequence had been produced in a study that used metagenomics to study parasite distribution in pig faeces [116]. The analysis shown in **Figure 13** was therefore repeated – this time without MK801450 (**Figure 14**).

Based on phylogenetic analysis of 562 nt positions (**Figure 14**), one might argue that a hypothesis of OP919601 representing a new subtype could be supported. However, in the analysis, the bootstrap value for ST1 could be higher, and therefore, similar to the situation accounted for in 2.3.1., the position of this sequence will remain unresolved until longer sequences are produced and/or more sequences accumulate that share 99%–100% similarity with OP919601. A somewhat similar situation is seen for MW718195; here, however, the deposited sequence is very short (259 bp).

Of note, two of the sequences in **Table 15** were deposited in GenBank as *Entamoeba polecki*, (AB845670, AB845671); however, these are not *Entamoeba*-specific sequences.

The geographical distribution of *E. polecki* is not well known, but positive samples have been observed from humans with recent traveling in Europe (Italy, Spain), Africa (Eritrea, Somalia, Ethiopia, and Kenya) and Asia (Afghanistan) [53]. Although there is a couple of reports on observations of *E. polecki* in humans in the Americas, no DNA data have yet been made available to corroborate the findings. The one DNA sequence from America available to date is from a pig (AF149913) and represents ST1.
It has not been possible to identify any cryptic host specificity within any of the subtypes so far based on SSU rDNA analysis; analysis of other genes appears relevant to investigate this further.

[2] Chimaeras arise during PCR amplification, usually when there are two distinct subtypes in the DNA sample and when there is incomplete replication of a DNA strand during a cycle [27].

2.3.7. *Entamoeba* in ungulates

The number of published studies on Entamoebas in ungulates that used molecular methods for detection and differentiation is relatively low. In 2010, we published data on *Entamoeba* in ruminants obtained by conventional PCR and Sanger sequencing [50]. Prior to the molecular work, cysts had been isolated from faecal samples by sucrose gradient centrifugation, and all cysts were uni-nucleate. Near-complete SSU rDNA sequences were obtained for six samples from cattle (n = 2), sheep (n = 2), reindeer (n = 1), and roe deer (n = 1). We argued that the sequences obtained from cattle, sheep and reindeer might be considered different genetic variants (genotypes) of *Entamoeba bovis*, and these data were the first DNA data made available for this species. Prior to our study, only morphological data were available for *E. bovis*. Meanwhile, the taxonomical status of the sequence obtained from the reindeer was unclear, as it might be considered a separate species based on the phylogenetic analysis.

There are currently 376 SSU rDNA sequences of *E. bovis* for which information on host species is available (data available on GitHub https://github.com/Entamoeba/DMSc-Thesis/blob/main/Entamoeba%20bovis.xlsx). Most sequences are from China (n = 290), but a few sequences are available from Australia, Brazil, Japan, and Sweden. The host species for *E. bovis* identified so far are summarised in **Table 16**. As seen, none of the sequences have been observed in non-artiodactyl hosts.

In our recent study of pigs [13], *Entamoeba polecki* was commonly seen, and, as expected, no subtypes other than ST1 and ST3 were observed. A minor proportion of the pigs (3%) were positive for *E. hartmanni*. No evidence of *E. suis* or other Entamoebas was obtained.

Our data could indicate that synanthropic and wild ungulates are commonly colonised by species of *Entamoeba*; however, of the hosts sampled to date, only pigs appear to be able to host Entamoebas that can colonise humans, namely *E. hartmanni* and *E. polecki*.

Quite a few other species of *Entamoeba* have been described based on samples from ungulates; however, in the absence of sequence data, the validity of species names such as *Entamoeba ovis*, *Entamoeba dilimani*, *Entamoeba bubalus*, *Entamoeba equibucalis*, *Entamoeba suigingivalis*, *Entamoeba gedoelsti*, and *Entamoeba caprae* remain unconfirmed.

2.4. ENDOLIMAX

Together with *Entamoeba* and *Iodamoeba*, *Endolimax* belongs to the Archamoebae, a group of anaerobic free-living or endobiotic protists that constitutes the major anaerobic lineage of the supergroup Amoebozoa [258]. The genus of *Endolimax* was described by Kuenen and Swellengrebel in 1917 [259]. Several species of *Endolimax* have been described in a range of host groups, including mammals, birds, reptiles, amphibians, and insects; for a comprehensive list of the species reported, see the review by Poulsen and Stensvold [31].

First described by Wenyon and O'Conner in 1917 [260], *Endolimax nana* is the smallest of the Archamoebae commonly infecting mankind, which is somehow reflected in its name. The mature cyst stage typically measures about 8–10 µm [229] and contains four nuclei that do not have peripheral chromatin, which make them relatively easy to recognise on light microscopy. Indeed, the parasite is usually detected by microscopy of faecal concentrates obtained by the traditional FECT. Specific primers have been developed and published by our groups for molecular detection and characterisation [31, 32, 192].

2.4.1. Aspects of biology, life cycle, diagnosis, and epidemiology of *Endolimax nana*

Positivity rates of about 30%–40% are not unusual in surveys carried out in some parts of the world, including Mexico [261], Colombia [262] and Brazil [263], but are typically somewhat lower. Meanwhile, in a survey of 3,374 children sampled in rural Côte d'Ivoire in the beginning of the millennium, 82.6% of the children tested positive for the parasite [264].
Based on a metanalysis, we recently estimated that 13.4% of the global gut-healthy population might be positive for *E. nana*, while only 3.4% of patients with GI symptoms may carry the parasite [31]. In Denmark, positivity rates of up to 7.5% have been seen, depending on study population [10, 90, 105, 189].

Using metabarcoding, we showed the presence of *E. nana*-specific DNA in 10/26 sewage samples from Sweden [33].

Apart from *E. nana*, *Endolimax piscium* is the only species of *Endolimax* for which SSU rDNA data is available to date (**Table 17**). *Endolimax piscium* was described in 2014 by Constenla et al. [265] as a cause of granulomatous disease in cultured Senegalese sole.

2.4.2 Genetic diversity and host specificity of *Endolimax nana*

Phylogenetically, *Endolimax* clusters closely together with *Iodamoeba* within the Mastigamoebidae B group of the Archamoebae. By combining PCR and TA cloning and Sanger sequencing and PacBio sequencing of pooled PCR products, we were recently able to describe the existence of two ribosomal lineages of *E. nana* [32] (**Figure 15**). These two lineages differ genetically by up to 16%. Surveys identifying the relative

positivity rates of these two lineages among *Endolimax*-positive individuals are still to be carried out with a view to identifying any differences in the epidemiology and clinical significance of the parasite.

It is clear from our phylogenetic analyses that *Endolimax* and *Iodamoeba* are sister taxa and cluster within Mastigamoebidae B, which confirms the findings of Pánek and colleagues [258]. However, despite the use of near-complete SSU rDNA sequences, it is still not possible to identify whether *E. piscium* is congeneric with *E. nana*. Near-complete SSUrDNA sequences from more species of *Endolimax* should be obtained and included in phylogenetic analysis to allow for more elaborate taxonomic inferences.

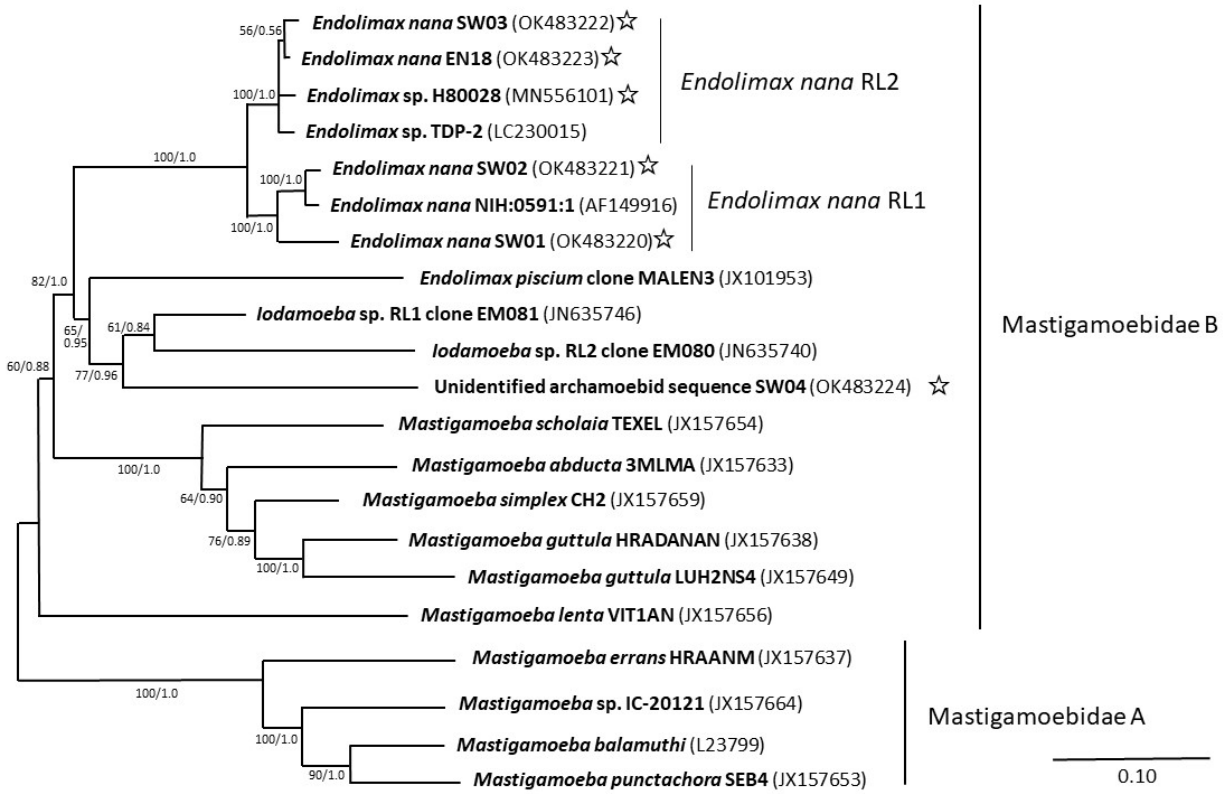


Figure 15. Maximum likelihood phylogeny of *Endolimax* and relatives (including *Iodamoeba* RL 1 and RL2), reconstructed from an SSU rDNA alignment consisting of 21 taxa and 2,067 positions. Maximum likelihood bootstrap values and Bayesian posterior probabilities are shown in that order on each bipartition. GenBank accession numbers are indicated in parentheses. The scale bar indicates nt substitutions per site. The sequences that were generated in our study [32] are indicated with a star; sequences from sewage have the prefix SW. Abbreviation: RL, ribosomal lineage. (Adapted from [32]).

As of the time of writing (January 2022), there are 34 *Endolimax*-specific sequences in the NCBI database from sewage, wastewater, stool, pig faeces, and fish muscle. The majority of these represent sequences of only ~100–700 bp, and these sequences are not all covering the same part of the SSU rRNA gene. Only 14 of the 34 sequences have more than 1,500 bp, but, again, they do not overlap completely in terms of gene coverage. Consequently, the data currently available for phylogenetic inferences are still very limited.

Based on publicly available DNA-sequence evidence (**Table 17**), two species have been identified as hosts of *E. nana*, namely *Homo sapiens* and *Sus scrofa domesticus* [266]. There are also reports of DNA-based detection of *E. nana* in non-human primates [267]. Microscopy studies have revealed high and moderate colonisation rates in macaques [268] and cercopithecids [269], but contrary to the situation for *Iodamoeba* (see below), no *E. nana*-specific sequences from non-human primates have yet made it into the NCBI database. More molecular studies are needed to delineate the host range for *E. nana*, but these preliminary data could indicate that the host range for both *E. nana* and *I. bütschlii* is quite similar, involving at least human and non-human primates as well as pigs.

2.5. IODAMOEBA

2.5.1. Aspects of biology, life cycle, diagnosis, and epidemiology of *Iodamoeba bütschlii*

Iodamoeba is a genus of intestinal parasitic protists found in humans, non-human primates, and other animals. The genus was described by Dobell in 1919 [195], who also gave the name *Iodamoeba bütschlii* to the human parasite; since then, *Iodamoeba* found in humans has been assigned to this species. Other species names introduced over the years include *Iodamoeba kuenenu*, *Iodamoeba suis*, and *Iodamoeba williamsi* [229]; however, it remains unknown, to which extent one or more of these are in fact *Iodamoeba bütschlii*.

The parasite has a typical Amoebozoan life cycle involving a cyst stage, which is the resting and infective stage, and a trophozoite stage, which is the feeding stage. Trophozoite stages may be difficult to tell apart from those of *Endolimax*.
The name '*Iodamoeba*' reflects the noticeable iodophilic glycogen mass present in *Iodamoeba* cysts often erroneously referred to as a vacuole. Cysts are irregularly shaped, vary in diameter with a mean of approximately 10 µm [195] and usually have a single vesicular nucleus with a large spherical karyosome and very little peripheral chromatin.

Among the genera of CLIPPs included in this thesis, *Iodamoeba* is probably among the least common in humans, if not *the* least common one. Positivity rates reported in surveys of GI parasites in humans sampled across the globe typically range between 0.1% and 2.0%. Of note, however, positivity rates between 7.5% and 15% were reported among 381 apparently healthy subjects from Camiri, Boyuibe, Gutierrez in Bolivia [270] and in surveys from Colombia [271], Venezuela [188], and Peru [272, 273], suggesting a relatively high transmission rate in the peri-equatorial part of South America. Observing *Iodamoeba* in a stool sample in the absence of other CLIPPs is a rare phenomenon (unpublished observations).

2.5.2. Genetic diversity of *Iodamoeba bütschlii*

It was only in 2012, that we published the first SSU rDNA sequences of *Iodamoeba* [29], and so *Iodamoeba* was the last genus of intestinal protozoa known to parasitise on humans to have its ribosomal DNA sequenced. The reason why this had taken so long has to do with the fact that sequencing *Iodamoeba*-specific DNA straight off PCR products with no post-PCR steps such as cloning is of limited use (see section 1 – Introduction) for this genus – a situation similar to that of *Endolimax*. The SSU rDNA sequences of *Iodamoeba* are longer than 2.5 kbp and thereby remarkably longer than those of for instance *Entamoeba*, *Dientamoeba* and *Blastocystis*, which are about 1.8 kbp.

Using DNA extracted from purified cysts confirmed as *Iodamoeba* by microscopy, low-specific primers, post-PCR cloning steps and sequence assembly techniques, we were successful in obtaining a few sequences reflecting *Iodamoeba*-specific SSU rRNA genes, six of which were longer than 1,000 bp; three sequences were even longer than 2,000 bp and thus covered what would be expected to be almost 80% of the gene.

We identified two ribosomal lineages (RL1 and RL2) that – given the remarkable genetic difference of ~30% – easily could be considered separate species – or even genera (**Figure 15**). The existence of these two lineages has since then been corroborated by independent studies [14, 33, 116], some of which involved metagenomics or metabarcoding. Moreover, there is potential evidence of yet another ribosomal lineage in the article by Hamad and colleagues [274], who sampled gorillas in Cameroon and deposited *Iodamoeba* sequences in GenBank (the JX- sequences in **Table 18**) that are all identical apart from two unique bp substitutions in two of the sequences. They appear to be more related to RL1 (94.5% similarity) than to RL2 (90.0% similarity) and may represent a new ribosomal lineage; however, these sequences do not align well with other *Iodamoeba* sequences, and no data on morphology were provided, so it is not known whether *Iodamoeba* cysts or trophozoites were in fact present in these samples. Not until longer sequences are provided (the sequences in question were only 697 bp long, representing less than a third of the gene), preferably along with morphology data from microscopy of faecal concentrates, it remains premature to conclude on this. The same situation applies to the seven sequences also referred to as "uncultured *Iodamoeba*" from cattle and sheep (**Table 18**), which data, however, were not accompanied by an article to my knowledge.

It should also be mentioned, that in our recent *Endolimax* study [32], we produced a sequence, 'SW04' (OK483224) **(Figure 15**) from one of the Swedish wastewater samples, which clustered with *Iodamoeba* RL1 and RL2 with relatively high bootstrap support, sitting on a long branch, which indicates a high degree of genetic divergence. Whether this sequence also represents a novel *Iodamoeba* lineage is uncertain as of yet; no morphology data were available for this sample either, which could have informed conclusions.

Together with its sister taxon, *Endolimax nana* (see above), *I. bütschlii* forms part of the Mastigamoebidae B group [176]. Apart from *Endolimax* and *Iodamoeba*, this group holds exclusively species of *Mastigamoeba*, which are mostly free-living amoeboflagellates.

2.5.3. Host specificity of *Iodamoeba bütschlii*

Apart from humans, *I. bütschlii* has been reported in non-human primates [29, 274-276] and pigs/wild boars [13, 29, 116, 276-278], for which hosts also cysts specific to *Iodamoeba* have been observed [276]. Moreover, there are sequences in the NCBI database that have been obtained from samples from sheep and cattle (**Table**

18), but these data would warrant confirmation. There is limited evidence that other natural hosts could be rodents, camels, and birds [29], but to my knowledge, there are no DNA sequences available for *Iodamoeba* from these hosts yet to confirm the findings.

Maybe not surprisingly, *I. bütschlii* (both RL1 and RL2) was found in the Swedish wastewater samples [33]. There are moreover two *Iodamoeba* sequences in the NCBI database isolated from waste water in Australia (MH623069 and MH623073; **Table 18**) belonging to RL1.

Positivity rates in NHP sampled in zoos range typically between 5% and 8% [279, 280]. However, a staggering 42.96% of 443 cynomolgus macaques bred in China and sampled in Italy tested positive in a relatively recent survey [268].

While both RL1 and RL2 have been found in humans, to date, only RL2 has been found in suid hosts. RL1 has been found in a macaque [29]; meanwhile, the many *Iodamoeba* sequences produced from samples from gorillas by Hamad and colleagues [274] may represent a new ribosomal lineage.

3. SUMMARY OF CLIPPs RARELY REPORTED IN HUMANS

A number of non-pathogenic Metamonad species have been reported only sporadically in surveys of humans (**Table 1**). Most of these organisms have life cycles that involve both a trophozoite and a cyst stage in their respective life cycles; meanwhile, *Pentatrichomonas* and *Trichomonas* appear to lack the cyst stage; a situation that may be similar to that of *Dientamoeba*.

The most common of these may be *Chilomastix mesnili*, which is reported with a frequency of at least up to 8.7% [270, 281, 282] in surveys from across the globe. With some reports for NHPs [268, 269, 283, 284], this species appears to be exclusive to primates. There are currently 12 SSU rDNA sequences in GenBank from two sources ([285] plus Čepička, unpublished [KC960584–KC960590]), which by alignment could appear to reflect at least two distinct subtypes – if not species, differing by ~9%; all sequences are from humans.

For *Enteromonas*, *Retortamonas* and *Pentatrichomonas*, survey data are so limited that it makes little sense to review the data or to speculate on the frequency by which these may show up in stool samples. However, in a recent survey of 127 faecal metagenomes from individuals sampled in Cameroon, Tanzania, Peru, Italy or USA, *Enteromonas*-specific DNA was identified in hunter-gatherer populations (6.2%–50%) and in Cameroonian fishers (15.8%), but not in individuals from Western countries [286]. The metagenomes included were not queried for *Retortamonas*- and *Pentatrichomonas*-specific data, and only a handful of samples displayed evidence of *C. mesnili*. To date, only three *Enteromonas* SSU rDNA sequences have been deposited in GenBank all of which were from the study by Kolisko and colleagues [287], and only one of which reflects *Enteromonas hominis* (EF551180); the remaining two sequences are from a turtle (EF551179) and a tortoise (EF551178), respectively. A prevalence of 2.4% was reported for *Retortamonas intestinalis* in the survey from Bolivia by Cancrini and colleagues [270], and in Libya, the positivity rate across 350 stool samples from children and neonates admitted to hospital was 3.43% [288]. Most of the 18S data in GenBank on *Retortamonas* stem from two sources: Hendarto and colleagues performed a study of vertebrate hosts (including humans), identifying a positivity rate of 4/290 (1.4%) in humans and 9/31 (29.0%) in water buffalos [289]; these hosts were the ones for which sampling was most extensive. The other larger set of SSU rDNA sequences were provided by Čepička in 2013. Of the 114 *Retortamonas*-specific DNA sequences currently available in GenBank, most are from humans, cattle, water buffalos, pigs, goats, rats and insects. Studies have agreed that *Retortamonas* from insects cluster separately from *Retortamonas* from vertebrate hosts. Interestingly, Retortamonads from vertebrates appear to cluster with diplomonads (which include *Giardia* and *Enteromonas*), whereas the ones from insects cluster with *Chilomastix* [289]. That Retortomonadida are polyphyletic have been confirmed by at least two research teams [289, 290].

With regards to *Pentatrichomonas*, survey data are also scarce, and some of the data published should be interpreted carefully, since reporting would rely on the detection of trophozoites in faecal material either by culture or by permanent staining methods, as this genus probably does not produce cysts, or amplification of specific DNA. Based on PCR, Li and colleagues identified a positivity rate of *Pentatrichomonas hominis* of 7.8% across 500 pig samples in China [291], and In Korea, Li and colleagues identified a positivity rate of 31.4% among 315 pet dogs [292]. A group of monkeys in China was also observed to be commonly colonised based on PCR data [293]. Meanwhile, positivity rates in humans in studies using reliable methods are up to 4.0% in northern China [293].

There are 180 18S sequences specific to *P. hominis* in GenBank with information on host source at the time of writing, with most sequences being from primates, including humans, and canine hosts. High positivity rates have been identified in *e.g.,* raccoon dogs [294]; however, it should be taken into consideration that surveys using nested PCR for detection might be at risk of overestimating the extent of factual colonisation rates due to issues with increased contamination risks and if specificity is not confirmed by Sanger sequencing. Of the genera mentioned in the present section, *Pentatrichomonas* is the one that is represented in GenBank with far most DNA sequences, including sequences reflecting coding genes. Analyses of nuclear SSU rRNA genes and flanking regions have not enabled distinction between *Pentatrichomonas* isolated from human and non-human hosts [295].

All local general and specific DNA-based attempts at SSI to screen for *Enteromonas* and *Pentatrichomonas* across samples from healthy individuals and individuals with GI symptoms in Denmark have failed to reveal any positive samples (Kaul and Stensvold, unpublished).

Trichomonas tenax is a parasite of the oral cavity of humans and other animals, especially dogs; other hosts include cat and horse [296]. It may co-colonise the oral cavity with *E. gingivalis* [231], and it may have a role in the development of periodontal disease [231, 236, 296]. Interestingly, a sequence named '*Trichomonas tenax*' (JX943581) is 99.73% similar to a sequence named '*Trichomonas canistomae*' (AY247748), and since other sequences with same coverage (99%) named '*Trichomonas tenax*' are less similar (*e.g.,* 99.38% for D49495), this could indicate that *T. tenax* and *T. canistomae* might be the same species. However, the data in GenBank for these organisms are scarce, so it is premature to do more in-depth analysis of this.

Conclusively, the organisms briefly and superficially reviewed in this section may not be as rare in humans as could be anticipated judged from available literature, as issues pertaining to lack of a faecal cyst stage (*Pentatrichomonas*), lack of reference data for developing primers for DNA-based screening and for mapping of DNA from *e.g.,* metagenomics studies (*Enteromonas* and *Retortamonas*) plus potentially large variation in prevalence according to geography and intrageneric diversity may have hampered attempts to develop a fuller picture of the colonisation rates in surveys of intestinal parasites.

4. CLIPPs AS MARKERS OF GASTROINTESTINAL HEALTH?

Based on what was then recent data from our own work and products of international collaboration [80, 90, 93, 94, 133, 297, 298], we published a comment in *Journal of Clinical Microbiology* asking the following question: '*Blastocystis* in health and disease: Are we moving from a clinical to a public health perspective?' [89].

This question embodied the acknowledgement of the fact that *Blastocystis* was generally more common in healthy individuals than in patients with irritable bowel syndrome, inflammatory bowel disease, and acute diarrhoea [90, 92, 298]. We had data suggesting that almost every third adult individual in Denmark is colonised with *Blastocystis* [90]. Although we had published two case reports describing successful eradication of *Blastocystis* followed by symptom resolution (one of which involved a rare subtype, ST8) [87, 88], we were unable to identify – based on a literature review – any medical treatment that could consistently lead to eradication of the parasite [83], and our general understanding is that *Blastocystis* colonisation might not typically be linked to the development of symptoms.

The healthy human gut is characterised by high alpha diversity and a predominance of obligate anaerobes [299, 300]. We hypothesised that *Blastocystis* might be linked to certain gut microbiota features, a hypothesis that was tested for the first time by our team using metagenomics data generated by the MetaHIT Consortium. We showed that *Blastocystis* was significantly more common in healthy individuals than in patients with Crohn's Disease, that it was linked to the *Ruminococcus* and *Prevotella* enterotypes, while inversely linked with the *Bacteroides* enterotype, and that there was a strong trend towards *Blastocystis* positivity being associated with low-to-normal body mass index [94]. Interestingly, the latter finding was later corroborated by a large metanalysis of 12 sets of metagenomics data by an independent research team [95].

Not only *Blastocystis*, but also *Entamoeba* has been linked to a lower relative abundance of Bacteroidetes and generally linked to eubiosis [301, 302], and the fact that the relative abundance of Proteobacteria was the phylum distinguishing *Blastocystis* carriers from non-carriers to the largest extent led us to speculate that also *Blastocystis* might generally be linked to eubiosis [45]. In our study led by Krogsgaard [93], we showed that bacterial alpha diversity of *Blastocystis* carriers was significantly higher than in non-carriers, a finding that was later corroborated using samples processed by metabarcoding in our own lab [45], and an observation that we also made for *Dientamoeba* carriers vs. non-carriers in a different study [303]. Fuelling the hypothesis I had stipulated a few years earlier [304], this experience and these findings led us question the clinical relevance of *Blastocystis*, but certainly prompted interest in continuing investigating *Blastocystis* and other CLIPPs as potential markers not only of gut health specifically, but also of overall public health.

The ‘Old Friends hypothesis’ embodies the theory that the development of the immune system relies on input from three sources, collectively referred to as the ‘old friends’: (i) the commensal microbiotas transmitted by the mother and other family members; (ii) organisms from the natural environment that modulate and diversify the commensal microbiotas; and (iii) the ‘old’ infections that could persist in small, isolated hunter-gatherer groups as relatively harmless subclinical infections or carrier states (colonisation). These categories of organisms have to be tolerated and hence play a role in the development and regulation of the immune system [305]. Comparing diversity patterns of intestinal eukaryotes between individuals with a westernised life style (USA) and individuals with an agrarian life cycle (Malawai), Parfrey and colleagues published data exemplifying the ‘defaunation’ of the human gut [306]. Interestingly, the authors noticed that individuals with non-western diets and life styles had microeukaryotic diversity patterns much more similar to those of non-human mammals compared with those with a western life style and diet. It has been hypothesised that the absence of exposure to parasites that used to commonly colonise and infect humans could result in the development of autoimmune diseases, such as IBD [307]. The hypothesis is indirectly supported by the fact that autoimmune diseases appear to be much more common in the Western world and it may also be a rare phenomenon in non-human mammals, for example. The question arises by which mechanisms parasites can mature the development of the host immune system. Some parasites may be in lifelong or lengthy direct contact with the immune system (*e.g., Toxoplasma gondii* and some parasitic nematodes), while others may not (*e.g.,* gut parasites such as CLIPPs); however, the latter may be influencing the immune system indirectly, perhaps by selecting for gut bacteria that do not interfere with the immune system in an inappropriate way.

The gut microbiota of patients with IBD differs from that of healthy individuals [308]. Patients with IBD experience a shift from strictly anaerobic bacteria towards facultative anaerobes such as the Enterobacteriaceae, indicating a role of oxygen in intestinal dysbiosis [309]. Reduced microbial diversity, increased Bacteroidetes and Enterobacteriaceae, and decreased Firmicutes proportions have all been observed in patients with IBD [310, 311], and the gut microbiota profiles of IBD patients to a large degree contrasts with that of *Blastocystis*-positive individuals. In our studies, we have generally found a lower prevalence of *Blastocystis* in patients with IBD [94, 126, 298]. It is likely that a parasite such as *Blastocystis* do not thrive in microaerophilic environments [96], and we have argued that colonisation by some intestinal parasites can be predicted with quite a high degree of accuracy merely by studying the composition of gut bacteria [96]. This might also indicate that *Blastocystis* may be merely an indicator organism rather than a gut microbiota manipulator.

Nevertheless, some micro-eukaryotes are known to exert a beneficial effects on the host. Parfrey and colleagues provided examples of mutualistic relationships between flagellates residing in the hindgut of termites and cockroaches where they break down cellulose [312]. These insects use parabasalid and oxymonad symbionts to break down cellulose and release energy, and these flagellates can constitute 15%–30% of the body weight of the termite. Looking at larger animals, ciliates are known to colonise the intestinal

tracts of a wide range of ruminant and non-ruminant herbivores. Although not essential for feed degradation and survival of the host, ciliates may contribute to overall gut function by adding degradative complexity, by their ability to scavenge oxygen, or by their grazing behaviour, which helps to shape and regulate prokaryotic populations [313]. Mishra and colleagues recently showed that the camel rumen eukaryotes (mainly ciliates) are highly dynamic and depend on the type of diet given to the animal [314], with two different types of feed selecting for two different types of micro-eukaryotes (Ciliophora vs. fungi); these observations could fuel investigations into CLIPPs as contributors to host metabolism and gut ecology homeostasis and potentially as markers of dietary intake of the host.

Over the past decade, attempts with faecal microbiota transplantation have been successful in terms of treating recurrent infection with *Clostridioides difficile* [315]. Since high microbiota diversity is an attractive asset of FMT donor stool, it is likely that a number of the donors used for obtaining FMT material are positive for *Blastocystis* and/or other CLIPPs. Although one study did not identify a difference in clinical improvement between recipients of *Blastocystis*-positive FMT material and those receiving *Blastocystis*-negative FMT material [316], the impact of the presence of CLIPPs in FMT donor stool should be investigated, and it should be explored to which extent CLIPPs can be 'transplanted' successfully from donor to recipient with FMT solutions.

5. DISCUSSION & FUTURE PERSPECTIVES

Before I started my interest in parasitology more than 20 years ago, very little was known regarding genetic diversity within *Entamoeba* species isolated from humans, and there was no evidence of cryptic genetic diversity in *E. nana* and *I. bütschlii*; in fact, there were no DNA data on *I. bütschlii* at all, a parasite colonising many millions of people as well as non-human primates and pigs, and there was only one DNA sequence of *E. nana*. Indeed, the whole field of molecular detection and differentiation of intestinal parsaites was in its very infancy. I introduced the idea of testing for *D. fragilis* in our laboratory, and we were the first ones to produce data on this parasite in Denmark; little did we know that this parasite would turn out to be a more or less obligate finding in children in Denmark. When I defended my PhD thesis on *Blastocystis* in 2008, only nine subtypes were known; no nuclear or mitochondrial genomic data had been published, and there was limited knowledge on host-specific differences within *Blastocystis* subtypes. Finally, we did not know that CLIPPs would typically be more common findings in gut-healthy individuals than in patients with GI symptoms, and, very importantly for the clinical aspect of CLIPPs, we had no idea that some of these parasites were markers of a healthy bacterial gut microbiota.

Prior to the introduction of DNA-based methods, it was customary to generate species names based on host range and morphology. We now know that both are unreliable, because host ranges can be broad, and identical morphology can hide substantial genetic differences. For some CLIPP genera, it may be so that we have only scratched the surface of universes of genetic diversity, the implications of which are still to be revealed (*e.g., Endolimax* and *Iodamoeba*); for one or two, we may have come a long way already (*e.g., Dientamoeba*).

A lot of the work included in this thesis (also including articles that were not shortlisted but were listed as supporting articles [p. 10–17]) has been central to informing our attempts to develop useful terminologies applicable by peers to similar data. Standardisation of terminologies is crucial to broaden the global understanding of the genetic diversity and host specificity of parasites in order to be able to delineate parasite transmission patterns and clinical and epidemiological differences. This was clear to us already in 2006, when we developed the subtype terminology for *Blastocystis* [26]. This terminology was quickly adapted by the community, and research into the distribution an overall epidemiology of *Blastocystis* subtypes has really taken off since then; however, we have had to keep an eye on the way new subtypes were introduced in order to reduce the risk of confusion [27]. Similar initiatives have been carried out for *Entamoeba* [181]. Given the complexity of the genetic diversity of some of these parasites, it is important that terminology is clear, robust and practical, and expert groups should revisit guidelines and standards on a continuous basis, the way it is currently done in the *Blastocystis* COST action (https://www.cost.eu/actions/CA21105/).

PCR coupled with Sanger sequencing for molecular characterisation and time-consuming and expensive primer walking for the mapping of *Blastocystis* mitochondrion-like genomes have been instrumental to the production of a lot of the data that has gone into this work. However, with more modern technologies, such as metabarcoding and Nanopore sequencing, data will be produced much faster and in greater quantity; longer reads can be produced with Nanopore sequencing, and we will have an opportunity to obtain a more comprehensive view of pro- and eukaryotic diversity when metabarcoding is used. We also have the opportunity to study links between parasites and other organisms, as exemplified in Section 4. In this process, however, one should take care to try and remember the relevance of including data on morphology where possible for reasons that should be clear from previous sections.

In terms of any use of metagenomics, it should also be obvious from the work included in this thesis that efficient use of metagenomics data relies on complete reference sequence databases. For many of the CLIPPs included in the present thesis only DNA sequences reflecting the SSU rRNA genes are available, and it might appear that not even these have yet been characterised fully. Data from WGS are still pending for most CLIPPs. At the time of writing, full genome data are available for five species of *Entamoeba* (*E. histolytica*, *E. dispar*, *E. moshkovskii*, *E. nuttalli*, and *E. invadens*), and for some subtypes of *Blastocystis*.

Even for our metabarcoding assay, there might be issues, despite the fact that this assay is targeting 18S genes only, which is probably the most commonly characterised gene across all living organism. The issue pertains to the fact that for each DNA sample processed by the assay, a varying number of sequence reads cannot be mapped to a reference taxon and therefore remain unannotated. Although these sequences can easily be extracted for each sample, the process of querying the sequences for new genetic variants is time consuming and of limited feasibility, with a rather limited the risk-reward trade-off, probably. Finally, this particular assay has varying degrees of sensitivity depending the parasite in question [13, 33], so in its current version it cannot be used as a one-fits-all (but maybe a one-fits-most) assay.

There is evidence that we only started to scratch the surface of eukaryotic diversity in complex matters such as faecal samples. Chouari et al. [317] used 18S sequencing to investigate eukaryotic diversity in wastewater, and of 1,519 analysed sequences, 160 operational taxonomic units (OTU) were identified. Altogether 56.9% of the phylotypes were assigned to novel phylogenetic molecular species, exhibiting <97% sequence similarity with their nearest affiliated representative within public databases. Similarly, Matsunaga et al. [318] observed that 60% of their 18S rRNA gene clones obtained from DNA extracted from municipal wastewater had <97% sequence identity to described eukaryotes. In both studies, data on *Blastocystis* and Amoebozoa were observed. These studies confirmed not only the vast DNA data gap in the eukaryotic tree of life, but also the relevance of using sewage as study material for investigations into eukaryotic diversity involving CLIPPs.

One of the most remarkable findings of our studies is the demonstration of vast cryptic genetic diversity in some of the species. For comparison, *E. dispar* only differs from *E. histolytica* by <2%, and from *E. bangladeshi* and *E. ecuadoriensis* by 5%–6% across SSU rRNA genes. Meanwhile, the genetic distance between *E. coli* subtypes of 12%–13% equals the distance between quadrinucleated *E. dispar* and uninucleated *E. bovis*, two species that clearly differ in both host spectrum and cyst morphology. Things become even more conspicuous when looking at the ribosomal lineages of *I. bütschlii*: The extent of genetic diversity across the two currently acknowledged lineages differs by as much as a good 30%. This figure might be difficult for us to grasp, when we remind ourselves that human and ovine or suid nuclear SSU rRNA genes differ only by a handful bp or two. The implications of this are yet not clear. It should be investigated whether this amount of genetic diversity is reflected in the remaining genome, in which case gene prediction analyses might provide us with an opportunity to study phenotypic differences among the lineages and whether there would be reason to hypothesise that the lineages differ in the impact on host ecology and host health overall. Single-cell genomics/transcriptomics may prove a way forward in this respect, potentially combined with the cyst isolation procedure that we used in the study of *Iodamoeba* [29], or on parasites in culture.

For some protozoa, at least two 'sets' of ribosomal genes exist, an asexual set and a sexual set; this was exemplified recently in our study on a case of *Plasmodium cynomologi* in a Danish tourist [319], where our metabarcoding assay picked up both types of SSU rDNA sequences. These two sets of sequences differ substantially. However, there is currently no evidence that something similar should exist within the Amoebozoa. Also, the evidence of links between lineages and geography would not support the hypothesis of two sets of ribosomal RNA genes in some of the Amoebozoa. Nevertheless, SSU rRNA genes are organised differently among the CLIPPs: In *Entamoeba*, rRNA genes are located exclusively on extrachromosomal plasmids (circular DNA) [320], whereas they are organised on different chromosomes in for instance *Blastocystis* (linear DNA) [54, 321]. This has implications for our interpretation of rRNA data and for what to expect in terms of intra-strain diversity.

Our research has shown that *Entamoeba* and *Blastocystis* are cosmopolitan parasites, possibly reaching every 'corner' of the world, including remote and frigid areas such as Greenland (unpublished observations). It is likely that the study of these parasites can further inform studies of the evolution and migration of host species (due to co-evolution). Indeed, it is also interesting that humans and several other larger species of mammals share some if not all of the CLIPP genera dealt with in this work and are indeed very common—in some instances almost obligatory—hosts; however, one major group of animals appears to be only accidental hosts of most of these, namely the carnivores. For instance, surveys of intestinal parasites in wild and synanthropic carnivores have revealed positivity rates going towards zero [106, 123, 322-325]. This may have to do with the theory that most Amoebozoa and *Blastocystis* are parasites lodged in the colon and maybe particularly in the caecum [229]. The latter anatomical structure is typically more or less lacking in animals that are predominantly carnivorous such as felines and canids. To this end, diet may play a major role here, since

herbi- and omnivores would eat relatively much fibre compared with carnivores, and the metabolism of fibre, which predominantly takes place in the caecum and adjacent parts of the intestine, results in the generation of short-chain fatty acids (SCFA) that are important for reducing oxygen levels and maintaining eubiosis in the gut [96]. SCFAs, which include butyrate and propionate, have several functions essential to colonic health and immune function and are known to regulate cells of both the innate and adaptive immune systems. With respect to intestinal homoeostasis, significant reductions in the abundance of bacteria involved in butyrate and propionate metabolism have been identified as markers of dysbiosis in ulcerative colitis [115]. It should also be explored to which extent CLIPPs would rely on SCFAs as a source of energy.
Another observation suggesting that carnivores are not natural hosts of these CLIPPs is that those few animals that do test positive for *Blastocystis,* for instance, appear not to be colonised by one or few select subtypes the way we usually see it for natural hosts (*e.g.,* suids being positive for typically ST1 and ST5, and bovids being positive for typically ST10 and ST14). The situation is similar for *e.g.* lemurs; contrary to NHPs, these are not common hosts of *Blastocystis*, but whenever *Blastocystis* would be found, no particular subtype would appear to predominate [51, 326]. This is an example of how molecular characterisation of parasites can assist in identifying natural and accidental hosts.

Contrary to cattle, pigs share a number of CLIPPs with primates. *Endolimax nana, E. hartmanni*, *E. polecki*, and *I. bütschlii* are all Amoebozoa shared by both pigs and primates. For *Blastocystis*, subtypes 1–3 can be seen in both pigs and primates, and ST5, which is particularly common in pigs appears common in apes, although not in human primates. Finally, *Dientamoeba*, a common coloniser of humans, has been observed in pigs by some research teams [327, 328]. The large overlap in micro-eukaryotic fauna may reflect the fact that pigs and humans are omnivorous while cattle are herbivorous and/or that pigs are genetically more related to primates than cattle (pigs and human share 98% of the DNA, while cattle and humans share about 80%), or at least that the gut microbiota of primates might be more similar to that of pigs than of any other non-primate host. A somewhat distant but still reminiscent scenario is seen for *Taenia*, where humans can serve as intermediate host of *Taenia solium* but not of *Taenia saginata*.

Taking the faecal-oral transmission for CLIPPs into account, it is striking that some of these – *Blastocystis* and *Dientamoeba* – are very common in a country such as Denmark, where hygiene practices are relatively high. If indeed pinworm is a suitable vector for *Dientamoeba*, then the relatively high occurrence of pinworm infection in Denmark (data not shown) might explain the common occurrence of *Dientamoeba* carriage in this country. For *Blastocystis*, the explanation may be less straightforward. However, given the fact that this organism belongs to a separate phylum – even kingdom – it might be of little use to assume that *Blastocystis* is limited to entirely the same way of transmission as cyst-forming protozoa. There may be characteristics of *Blastocystis* that enable it to exist and persist in the environment to an extent that we are not aware of and that enables the organism to be transmitted in a way that differs from what is seen for members of the Archamoebae. By all means, the colonisation pressure of *Blastocystis* must be enormous.

When evaluating observations from DNA-based surveys of species/ribosomal lineages of CLIPPs there are a couple of things to bear in mind. Ideally, full-length SSU rDNA sequences should be obtained, but to date, most researchers have used 'partial sequences'; *i.e.,* sequences reflecting only part of the SSU rRNA gene, typically sequence fragments that could be covered by conventional bidirectional Sanger sequencing (up to about 800 bp). Within the field, there have only been few attempts towards advocating for a standardisation of primers used for amplification and sequencing, and so there are numerous examples of differences in the SSU rRNA gene regions covered. For example, Calegar and colleagues [198] used the *Entamoeba* primers developed by Verweij et al. [257] , which cover a region different to that explored by our group, namely the 5'-end of the SSU rRNA gene. Meanwhile, in our recent study on the diversity of *E. coli* and *E. hartmanni*, we used low-specificity primers to amplify the middle section of the *E. hartmanni* SSU rRNA gene [13]. This part of the gene apparently holds more genetic information and enable better resolution than the 5'-end of the gene. It should also be investigated whether the 'Entam' primers developed more than 20 years ago (*i.e.,* at a time when the NCBI database held a very limited number of *Entamoeba*-specific DNA sequences) for genus-specific amplification are oligos targeting DNA sequences that are conserved among all species of *Entamoeba*; at least for the reverse primer ('Entam2'), there is a single mismatch compared with *E. suis* (DQ286372) and *E. gingivalis* DNA sequences (KX027297, D28490). At least one mismatch to most if not all Entamoebas is also seen in the 'TN14' reverse primer developed by Matey et al. [193], which was also recently used in the study by Mulinge and colleagues [329]. The genus-specific primers published by our group in 2011 ('ENTAGEN_F' and 'ENTAGEN_R' [49]) still appear to be useful as we have not been able to identify sequence variation in *Entamoeba* in the primer annealing regions. Also for *Blastocystis*, various regions of the SSU rRNA gene have been studied (*e.g.,* 'Scicluna region' [46], 'Santin region' [330], and 'Stensvold region' [66]).

Nested PCR approaches have been used numerous times to detect and differentiate CLIPPs. Sometimes, the inner PCR has been designed as a species-specific PCR, and the advantage of that would be that the work and expense of sequencing could be obviated, since the species diagnosis would be carried out based on the size of the PCR product alone or maybe by RFLP. However, this methodology is also associated with drawbacks. The primers will only amplify what is already known, and any new variants of the species in question might therefore not be detected, in case this variation exists in the primer annealing region. Secondly, if PCR products are not sequenced, false-positivity is a possibility, depending on the quality of assay validation. Thirdly, in a study yet unpublished, DNA sequences of *E. hartmanni* were deposited in 2022 in the NCBI database (OP688358-OP688362), with the domestic dog listed as the host. In this very instance, sequence data are in fact available. Nevertheless, it should also be noted that if this data is a result of nested PCR, the very significance of the findings should be interpreted with caution, given the fact that dogs are coprophagous. Parasites may be able to pass through the digestive tract of hosts that are not natural hosts and still be detected by highly sensitive DNA-based methods such as nested PCR in faeces even if present in only very small

numbers. PCR does not distinguish between DNA from live and dead organisms. The best way to clearly demonstrate dogs as 'new' hosts of *E. hartmanni* would therefore be to establish the parasite in culture of a dog faecal sample (which would strongly suggest a live isolate that had in fact been colonising the dog's intestine), or at least produce a sequence from only a single-round PCR, which would typically be a less sensitive method than nested PCR.

While at the time of writing, it still appears likely that CLIPPs *per se* do not inflict disease on humans to any major extent, the roles of CLIPPs as 'Trojan horses' of for instance viruses should be explored further. It has been known for long that parasites can host bacteria (*e.g., Acanthamoeba* can host *Legionella* [331]), and moreover, there is recent evidence suggesting many CLIPP species as hosts for pecoviruses, hudisaviruses, Kirkoviridae, and Redondoviridae, among others [232, 332]. Wider use of metagenomics is expected to enable the disclosure of such relationships.

DNA from intestinal parasites that are common and that exhibit high degrees of host specificity can be used as strong indicative evidence of host in cases where faecal material is available for analysis by *e.g.,* metabarcoding, and where DNA data are insufficient for host identification (humans and some other larger mammals differ very little across the 18S gene).

In clinical microbiology laboratories the reporting of CLIPPs has been 'good practice', also in settings where the general consensus has been not to treat. The reporting of CLIPPs has been relevant especially to raise awareness of faecal-oral exposure, which could prompt further investigations for intestinal pathogens. The problem here is that colonisation by CLIPPs can be lengthy [80], and, unless there is a recent reference sample that was negative for CLIPPs, the reporting of CLIPPs in stool may have limited value. Paraclinical findings, such as Charcot-Leyden crystals (break-down products of eosinophils and basophils) and/or blood cells in stool, which may be revealed by microscopy of faecal concentrates, or better, by direct microscopy or permanent staining of fixed faeces, might be more relevant information; however, the practice of reporting these findings is probably declining.

Summary of outstanding questions and activities:

1. Update reference DNA sequence databases with genomic data (near-complete 18S sequences or even entire ribosomal operons, genomes of organelles where possible, and nuclear genomes).
2. Try to interpret what the extreme genetic diversity seen in *Iodamoeba* and *Endolimax* tells us and investigate whether the extensive genetic diversity observed across their SSU rRNA genes is reflected elsewhere in their genomes. What impact does their level of intrageneric diversity have on exisiting species concept(s)?
3. Expand on the knowledge of the genetic diversity and the host spectrum of CLIPPs by sampling more and different hosts.

4. Expand on the knowledge of the genetic diversity and the host spectrum of CLIPPs by sampling more and different hosts.
5. Obtain “helicopter views” of microeukaryotic diversity in human and non-human faecal samples using metagenomics and OTU annotation together with bioinformatics tools to identify hitherto unexplored diversity. Findings could be coupled with metadata and used to identify OTU communities linked to demographic features, life styles (including diets), hosts, diseases, *etc.* OTUs of specific interest could be characterised by rDNA full-operon analysis using *e.g.,* Nanopore sequencing.
6. Explore the metabolism of CLIPPs by genomic *in-vitro* predictions or by wet lab experiments, including the ability of CLIPPs to ‘predate’ on the host gut microbiome, including investigations into the enzymes and metabolites released by CLIPPs.
7. Explore the role of CLIPP colonisation on host microbiomes. Animal or advanced *in-vitro* models that can mimic *in-vivo* models are warranted. It should also be investigated for how long one can be colonised with species of Amoebozoa.
8. Explore the role of CLIPPs as hosts and transmitters of bacteria and viruses.

6. REFERENCES

1. Thomas-Lopez D, Muller L, Vestergaard LS, *et al.* Veterinary students have a higher risk of contracting cryptosporidiosis when calves with high fecal *Cryptosporidium* loads are used for fetotomy exercises. Appl Environ Microbiol 2020; 86(19): e01250-20.
2. Michlmayr D, Alves de Sousa L, Müller L, *et al.* Incubation period, spore shedding duration, and symptoms of *Enterocytozoon bieneusi* Genotype C infection in a foodborne outbreak in Denmark, 2020. Clin Infect Dis 2022; 75(3): 468-75.
3. Ethelberg S, Lisby M, Vestergaard LS, *et al.* A foodborne outbreak of *Cryptosporidium hominis* infection. Epidemiol Infect 2009; 137(3): 348-56.
4. Lebbad M, Winiecka-Krusnell J, Stensvold CR, Beser J. High diversity of *Cryptosporidium* species and subtypes identified in cryptosporidiosis acquired in Sweden and abroad. Pathogens 2021; 10(5): 523.
5. Widerström M, Schönning C, Lilja M, *et al.* Large outbreak of *Cryptosporidium hominis* infection transmitted through the public water supply, Sweden. Emerg Infect Dis 2014; 20(4): 581-9.
6. Decraene V, Lebbad M, Botero-Kleiven S, Gustavsson AM, Löfdahl M. First reported foodborne outbreak associated with microsporidia, Sweden, October 2009. Epidemiol Infect 2012; 140(3): 519-27.
7. Robertson LJ, Temesgen TT, Tysnes KR, Eikås JE. An apple a day: an outbreak of cryptosporidiosis in Norway associated with self-pressed apple juice. Epidemiol Infect 2019; 147: e139.
8. Hanevik K, Hausken T, Morken MH, et al. Persisting symptoms and duodenal inflammation related to *Giardia duodenalis* infection. J Infect 2007; 55(6): 524-30.
9. Stensvold CR. Pinning down the role of common luminal intestinal parasitic protists in human health and disease - status and challenges. Parasitology 2019; 146(6): 695-701.
10. Stensvold CR, Nielsen HV. Comparison of microscopy and PCR for detection of intestinal parasites in Danish patients supports an incentive for molecular screening platforms. J Clin Microbiol 2012; 50(2): 540-1.
11. Verweij JJ, Stensvold CR. Molecular testing for clinical diagnosis and epidemiological investigations of intestinal parasitic infections. Clin Microbiol Rev 2014; 27(2): 371-418.
12. Stensvold CR, Lebbad M, Verweij JJ. The impact of genetic diversity in protozoa on molecular diagnostics. Trends Parasitol 2011; 27(2): 53-8.
13. Stensvold CR, Jirku-Pomajbikova K, Tams KW, *et al.* Parasitic intestinal protists of zoonotic relevance detected in pigs by metabarcoding and real-time PCR. Microorganisms 2021; 9(6): 1189.
14. Chihi A, O'Brien Andersen L, Aoun K, Bouratbine A, Stensvold CR. Amplicon-based next-generation sequencing of eukaryotic nuclear ribosomal genes (metabarcoding) for the detection of single-celled parasites in human faecal samples. Parasite Epidemiol Control 2022; 17: e00242.
15. Holmgaard DB, Barnadas C, Mirbarati SH, Andersen LO, Nielsen HV, Stensvold CR. Detection and identification of *Acanthamoeba* and other non-viral causes of infectious keratitis in corneal scrapings by real-time PCR and next-generation sequencing-based 16S-18S gene analysis. J Clin Microbiol 2020; 59(2): e02224-20.
16. Stensvold CR, Lebbad M, Verweij JJ, *et al.* Identification and delineation of members of the *Entamoeba* complex by pyrosequencing. Mol Cell Probes 2010; 24(6): 403-6.
17. Andersen LO, Vedel Nielsen H, Stensvold CR. Waiting for the human intestinal Eukaryotome. ISME J 2013; 7(7): 1253-5.
18. Stensvold CR, Clark CG. Current status of *Blastocystis*: A personal view. Parasitol Int 2016; 65(6 Pt B): 763-71.
19. Clark CG. Cryptic genetic variation in parasitic protozoa. J Med Microbiol 2000; 49(6): 489-91.

20. Diamond LS, Clark CG. A redescription of *Entamoeba histolytica* Schaudinn, 1903 (Emended Walker, 1911) separating it from *Entamoeba dispar* Brumpt, 1925. J Eukaryot Microbiol 1993; 40(3): 340-4.
21. Clark CG, Diamond LS. Ribosomal RNA genes of 'pathogenic' and 'nonpathogenic' *Entamoeba histolytica* are distinct. Mol Biochem Parasitol 1991; 49(2): 297-302.
22. Kumanan T, Sujanitha V, Sreeharan N. Amoebic liver abscess: a neglected tropical disease. Lancet Infect Dis 2020; 20(2): 160-2.
23. Shirley DA, Moonah S. Fulminant amebic colitis after corticosteroid therapy: a systematic review. PLoS Negl Trop Dis 2016; 10(7): e0004879.
24. Ralston KS, Petri WA. Tissue destruction and invasion by *Entamoeba histolytica*. Trends Parasitol 2011; 27(6): 254-63.
25. Novati S, Sironi M, Granata S, *et al.* Direct sequencing of the PCR amplified SSU rRNA gene of *Entamoeba dispar* and the design of primers for rapid differentiation from *Entamoeba histolytica*. Parasitology 1996; 112 (Pt 4): 363-9.
26. Stensvold CR, Suresh GK, Tan KS, *et al.* Terminology for *Blastocystis* subtypes–a consensus. Trends Parasitol 2007; 23(3): 93-6.
27. Stensvold CR, Clark CG. Pre-empting Pandora's box: *Blastocystis* subtypes revisited. Trends Parasitol 2020; 36(3): 229-32.
28. Jacob AS, Busby EJ, Levy AD, Komm N, Clark CG. Expanding the *Entamoeba* universe: new hosts yield novel ribosomal lineages. J Eukaryot Microbiol 2016; 63(1): 69-78.
29. Stensvold CR, Lebbad M, Clark CG. Last of the human protists: the phylogeny and genetic diversity of *Iodamoeba*. Mol Biol Evol 2012; 29(1): 39-42.
30. Silberman JD, Clark CG, Diamond LS, Sogin ML. Phylogeny of the genera *Entamoeba* and *Endolimax* as deduced from small-subunit ribosomal RNA sequences. Mol Biol Evol 1999; 16(12): 1740-51.
31. Poulsen CS, Stensvold CR. Systematic review on *Endolimax nana:* a less well studied intestinal ameba. Trop Parasitol 2016; 6(1): 8-29.
32. Hocke EF, Jamy M, Burki F, Clark CG, Stensvold CR. Unravelling the phylogeny of a common intestinal protist: intrageneric diversity of *Endolimax*. Protist 2022; 173(5): 125908.
33. Stensvold CR, Lebbad M, Hansen A, *et al.* Differentiation of *Blastocystis* and parasitic archamoebids encountered in untreated wastewater samples by amplicon-based next-generation sequencing. Parasite Epidemiol Control 2020; 9: e00131.
34. Scanlan PD, Marchesi JR. Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and -independent analysis of faeces. ISME J 2008; 2(12): 1183-93.
35. Lukeš J, Stensvold CR, Jirků-Pomajbíková K, Wegener Parfrey L. Are human intestinal eukaryotes beneficial or commensals? PLoS Pathog 2015; 11(8): e1005039.
36. Fredensborg BL, Fossdal IKI, Johannesen TB, Stensvold CR, Nielsen HV, Kapel CMO. Parasites modulate the gut-microbiome in insects: A proof-of-concept study. PLoS One 2020; 15(1): e0227561.
37. Růžková J, Květoňová D, Jirků M, *et al.* Evaluating rodent experimental models for studies of *Blastocystis* ST1. Exp Parasitol 2018; 191: 55-61.
38. Kumar S, Stecher G, Tamura K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 2016; 33(7): 1870-4.
39. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 2013; 30(12): 2725-9.
40. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 2011; 28(10): 2731-9.
41. Stensvold CR, Alfellani M, Clark CG. Levels of genetic diversity vary dramatically between *Blastocystis* subtypes. Infect Genet Evol 2012; 12(2): 263-73.
42. Staden R, Beal KF, Bonfield JK. The Staden package, 1998. Methods Mol Biol 2000; 132: 115-30.

43. Stensvold CR, Nielsen M, Baraka V, Lood R, Fuursted K, Nielsen HV. *Entamoeba gingivalis:* epidemiology, genetic diversity and association with oral microbiota signatures in North Eastern Tanzania. J Oral Microbiol 2021; 13(1): 1924598.
44. Stensvold CR, Ascuña-Durand K, Chihi A, *et al.* Further insight into the genetic diversity of *Entamoeba coli* and *Entamoeba hartmanni*. J Eukaryot Microbiol 2022: e12949.
45. Stensvold CR, Sorland BA, Berg R, *et al.* Stool microbiota diversity analysis of *Blastocystis*-positive and *Blastocystis*-negative individuals. Microorganisms 2022; 10(2): 326.
46. Scicluna SM, Tawari B, Clark CG. DNA barcoding of *Blastocystis*. Protist 2006; 157(1): 77-85.
47. Stensvold CR, Clark CG. Molecular identification and subtype analysis of *Blastocystis*. Curr Protoc Microbiol 2016; 43: 20A.2.1-A.2.10.
48. Stensvold CR. Comparison of sequencing (barcode region) and sequence-tagged-site PCR for *Blastocystis* subtyping. J Clin Microbiol 2013; 51(1): 190-4.
49. Stensvold CR, Lebbad M, Victory EL, *et al.* Increased sampling reveals novel lineages of *Entamoeba:* consequences of genetic diversity and host specificity for taxonomy and molecular detection. Protist 2011; 162(3): 525-41.
50. Stensvold CR, Lebbad M, Clark CG. Genetic characterisation of uninucleated cyst-producing *Entamoeba* spp. from ruminants. Int J Parasitol 2010; 40(7): 775-8.
51. Stensvold CR, Alfellani MA, Nørskov-Lauritsen S, *et al.* Subtype distribution of *Blastocystis* isolates from synanthropic and zoo animals and identification of a new subtype. Int J Parasitol 2009; 39(4): 473-9.
52. Stensvold CR, Clark CG, Röser D. Limited intra-genetic diversity in *Dientamoeba fragilis* housekeeping genes. Infect Genet Evol 2013; 18: 284-6.
53. Stensvold CR, Winiecka-Krusnell J, Lier T, Lebbad M. Evaluation of a PCR method for detection of *Entamoeba polecki*, with an overview of its molecular epidemiology. J Clin Microbiol 2018; 56(5): e00154-18.
54. Hoevers JD, Snowden KF. Analysis of the ITS region and partial ssu and lsu rRNA genes of *Blastocystis* and *Proteromonas lacertae*. Parasitology 2005; 131(Pt 2): 187-96.
55. Kostka M, Hampl V, Cepicka I, Flegr J. Phylogenetic position of *Protoopalina intestinalis* based on SSU rRNA gene sequence. Mol Phylogenet Evol 2004; 33(1): 220-4.
56. El Safadi D, Gaayeb L, Meloni D, *et al.* Children of Senegal River Basin show the highest prevalence of *Blastocystis* sp. ever observed worldwide. BMC Infect Dis 2014; 14: 164.
57. Poulsen CS, Efunshile AM, Nelson JA, Stensvold CR. Epidemiological aspects of *Blastocystis* colonization in children in Ilero, Nigeria. Am J Trop Med Hyg 2016; 95(1): 175-9.
58. Aggarwal D, Mitra S, Kalra P, Bagga B, Mishra D, Takkar B. Case report: poor outcome despite aggressive management in *Pythium insidiosum* endophthalmitis. Am J Trop Med Hyg 2022; 108(1): 27-30.
59. Stensvold CR, Tan KSW, Clark CG. *Blastocystis*. Trends Parasitol 2020; 36(3): 315-6.
60. Rene BA, Stensvold CR, Badsberg JH, Nielsen HV. Subtype analysis of *Blastocystis* isolates from *Blastocystis* cyst excreting patients. Am J Trop Med Hyg 2009; 80(4): 588-92.
61. Clark CG, Stensvold CR. *Blastocystis:* Isolation, xenic cultivation, and cryopreservation. Curr Protoc Microbiol 2016; 43: 20A.1.1-A.1.8.
62. Stensvold CR, Arendrup MC, Jespersgaard C, Mølbak K, Nielsen HV. Detecting *Blastocystis* using parasitologic and DNA-based methods: a comparative study. Diagn Microbiol Infect Dis 2007; 59(3): 303-7.
63. Stensvold CR. Laboratory diagnosis of *Blastocystis* spp. Trop Parasitol 2015; 5(1): 3-5.
64. Süli T, Kozoderović G, Potkonjak A, *et al.* Subtyping *Blastocystis* in pigs and humans revealed unusual avian-specific subtype ST6 in humans in Serbia. Zoonoses Public Health 2021; 68(5): 544-8.
65. Vaisusuk K, Saijuntha W, Sedlak S, *et al. Blastocystis* subtypes detected in long-tailed macaques in Thailand - further evidence of cryptic host specificity. Acta Trop 2018; 184: 78-82.

66. Stensvold R, Brillowska-Dabrowska A, Nielsen HV, Arendrup MC. Detection of *Blastocystis hominis* in unpreserved stool specimens by using polymerase chain reaction. J Parasitol 2006; 92(5): 1081-7.
67. Stensvold CR, Ahmed UN, Andersen LO, Nielsen HV. Development and evaluation of a genus-specific, probe-based, internal process controlled real-time PCR assay for sensitive and specific detection of *Blastocystis*. J Clin Microbiol 2012; 50(6): 1847-51.
68. Nourrisson C, Brunet J, Flori P, *et al.* Comparison of DNA extraction methods and real-time PCR assays for the detection of *Blastocystis* sp. in stool specimens. Microorganisms 2020; 8(11): 1768.
69. Pérez-Brocal V, Clark CG. Analysis of two genomes from the mitochondrion-like organelle of the intestinal parasite *Blastocystis:* complete sequences, gene content, and genome organization. Mol Biol Evol 2008; 25(11): 2475-82.
70. Wawrzyniak I, Roussel M, Diogon M, *et al.* Complete circular DNA in the mitochondria-like organelles of *Blastocystis hominis.* Int J Parasitol 2008; 38(12): 1377-82.
71. Jacob AS, Andersen LO, Bitar PP, *et al. Blastocystis* mitochondrial genomes appear to show multiple independent gains and losses of start and stop codons. Genome Biol Evol 2016; 8(11): 3340-50.
72. Denoeud F, Roussel M, Noel B, *et al.* Genome sequence of the stramenopile *Blastocystis*, a human anaerobic parasite. Genome Biol 2011; 12(3): R29.
73. Clark CG, van der Giezen M, Alfellani MA, Stensvold CR. Recent developments in *Blastocystis* research. Adv Parasitol 2013; 82: 1-32.
74. Jinatham V, Maxamhud S, Popluechai S, Tsaousis AD, Gentekaki E. *Blastocystis* one health approach in a rural community of northern Thailand: prevalence, subtypes and novel transmission routes. Front Microbiol 2021; 12: 746340.
75. Jinatham V, Nonebudsri C, Wandee T, Popluechai S, Tsaousis AD, Gentekaki E. *Blastocystis* in tap water of a community in northern Thailand. Parasitol Int 2022; 91: 102624.
76. Martín-Escolano R, Ng GC, Tan KSW, Stensvold CR, Gentekaki E, Tsaousis AD. Resistance of *Blastocystis* to chlorine and hydrogen peroxide. Parasitol Res 2023; 122(1): 167-76.
77. Cabrine-Santos M, Cintra EoN, do Carmo RA, *et al.* Occurrence of *Blastocystis* spp. In Uberaba, Minas Gerais, Brazil. Rev Inst Med Trop Sao Paulo 2015; 57(3): 211-4.
78. Abdulsalam AM, Ithoi I, Al-Mekhlafi HM, *et al.* Prevalence, predictors and clinical significance of *Blastocystis* sp. in Sebha, Libya. Parasit Vectors 2013; 6: 86.
79. Stensvold CR, Arendrup MC, Nielsen HV, Mølbak K. [*Blastocystis*–an enigmatic parasite]. Ugeskr Laeger 2009; 171(34): 2388-90.
80. Scanlan PD, Stensvold CR, Rajilić-Stojanović M, *et al.* The microbial eukaryote *Blastocystis* is a prevalent and diverse member of the healthy human gut microbiota. FEMS Microbiol Ecol 2014; 90(1): 326-30.
81. Stensvold CR, Nielsen HV, Mølbak K, Smith HV. Pursuing the clinical significance of *Blastocystis* – diagnostic limitations. Trends Parasitol 2009; 25(1): 23-9.
82. Coyle CM, Varughese J, Weiss LM, Tanowitz HB. *Blastocystis:* to treat or not to treat. Clin Infect Dis 2012; 54(1): 105-10.
83. Stensvold CR, Smith HV, Nagel R, Olsen KE, Traub RJ. Eradication of *Blastocystis* carriage with antimicrobials: reality or delusion? J Clin Gastroenterol 2010; 44(2): 85-90.
84. Engsbro AL, Stensvold CR. *Blastocystis:* to treat or not to treat...but how? Clin Infect Dis 2012; 55(10): 1431-2.
85. Bremer Christensen C, Soelberg J, Stensvold CR, Jäger AK. Activity of medicinal plants from Ghana against the parasitic gut protist *Blastocystis*. J Ethnopharmacol 2015; 174: 569-75.
86. Nagel R, Cuttell L, Stensvold CR, Mills PC, Bielefeldt-Ohmann H, Traub RJ. *Blastocystis* subtypes in symptomatic and asymptomatic family members and pets and response to therapy. Intern Med J 2012; 42(11): 1187-95.
87. Stensvold CR, Arendrup MC, Nielsen HV, Bada A, Thorsen S. Symptomatic infection with *Blastocystis* sp. subtype 8 successfully treated with trimethoprim-sulfamethoxazole. Ann Trop Med Parasitol 2008; 102(3): 271-4.

88. Vogelberg C, Stensvold CR, Monecke S, *et al. Blastocystis* sp. subtype 2 detection during recurrence of gastrointestinal and urticarial symptoms. Parasitol Int 2010; 59(3): 469-71.
89. Andersen LO, Stensvold CR. *Blastocystis* in health and disease: are we moving from a clinical to a public health perspective? J Clin Microbiol 2016; 54(3): 524-8.
90. Krogsgaard LR, Engsbro AL, Stensvold CR, Nielsen HV, Bytzer P. The prevalence of intestinal parasites is not greater among individuals with irritable bowel syndrome: a population-based case-control study. Clin Gastroenterol Hepatol 2015; 13(3): 507-13.e2.
91. Petersen AM, Nielsen HV, Stensvold CR, *et al. Blastocystis* and *Dientamoeba fragilis* in active and inactive inflammatory bowel disease. Gastroenterology 2011; 140(5): S329-S30.
92. Stensvold CR, Christiansen DB, Olsen KE, Nielsen HV. *Blastocystis* sp. subtype 4 is common in Danish *Blastocystis*-positive patients presenting with acute diarrhea. Am J Trop Med Hyg 2011; 84(6): 883-5.
93. Krogsgaard LR, Andersen LO, Johannesen TB, *et al.* Characteristics of the bacterial microbiome in association with common intestinal parasites in irritable bowel syndrome. Clin Transl Gastroenterol 2018; 9(6): 161.
94. Andersen LO, Bonde I, Nielsen HB, Stensvold CR. A retrospective metagenomics approach to studying *Blastocystis*. FEMS Microbiol Ecol 2015; 91(7): fiv072.
95. Beghini F, Pasolli E, Truong TD, Putignani L, Cacciò SM, Segata N. Large-scale comparative metagenomics of *Blastocystis*, a common member of the human gut microbiome. ISME J 2017; 11(12): 2848-63.
96. Stensvold CR, van der Giezen M. Associations between gut microbiota and common luminal intestinal parasites. Trends Parasitol 2018; 34(5): 369-77.
97. Yoshikawa H, Nagano I, Wu Z, Yap EH, Singh M, Takahashi Y. Genomic polymorphism among *Blastocystis hominis* strains and development of subtype-specific diagnostic primers. Mol Cell Probes 1998; 12(3): 153-9.
98. Parkar U, Traub RJ, Vitali S, *et al.* Molecular characterization of *Blastocystis* isolates from zoo animals and their animal-keepers. Vet Parasitol 2010; 169(1-2): 8-17.
99. Fayer R, Santin M, Macarisin D. Detection of concurrent infection of dairy cattle with *Blastocystis*, *Cryptosporidium*, *Giardia*, and *Enterocytozoon* by molecular and microscopic methods. Parasitol Res 2012; 111(3): 1349-55.
100. Alfellani MA, Taner-Mulla D, Jacob AS, *et al.* Genetic diversity of *Blastocystis* in livestock and zoo animals. Protist 2013; 164(4): 497-509.
101. Su C, Mei X, Feng X, *et al.* Prevalence and molecular subtyping of *Blastocystis* sp. in rabbits in Henan, Central China. Folia Parasitol (Praha) 2022; 69: 2022.027.
102. Baek S, Maloney JG, Molokin A, George NS, Cortés Vecino JA, Santin M. Diversity of *Blastocystis* subtypes in horses in Colombia and identification of two new subtypes. Microorganisms 2022; 10(9): 1693.
103. Maloney JG, Molokin A, Santin M. Use of Oxford Nanopore MinION to generate full-length sequences of the *Blastocystis* small subunit (SSU) rRNA gene. Parasit Vectors 2020; 13(1): 595.
104. Alfellani MA, Stensvold CR, Vidal-Lapiedra A, Onuoha ES, Fagbenro-Beyioku AF, Clark CG. Variable geographic distribution of *Blastocystis* subtypes and its potential implications. Acta Trop 2013; 126(1): 11-8.
105. Stensvold CR, Nielsen SD, Badsberg JH, *et al.* The prevalence and clinical significance of intestinal parasites in HIV-infected patients in Denmark. Scand J Infect Dis 2011; 43(2): 129-35.
106. Paulos S, Köster PC, de Lucio A, *et al.* Occurrence and subtype distribution of *Blastocystis* sp. in humans, dogs and cats sharing household in northern Spain and assessment of zoonotic transmission risk. Zoonoses Public Health 2018; 65(8): 993-1002.
107. Forsell J, Granlund M, Stensvold CR, Clark CG, Clark GC, Evengård B. Subtype analysis of *Blastocystis* isolates in Swedish patients. Eur J Clin Microbiol Infect Dis 2012; 31(7): 1689-96.

108. Lhotska Z, Jirku M, Hlozkova O, *et al.* A study on the prevalence and subtype diversity of the intestinal protist *Blastocystis* sp. in a gut-healthy human population in the Czech Republic. Front Cell Infect Microbiol 2020; 10: 544335.
109. Stensvold CR. *Blastocystis:* Genetic diversity and molecular methods for diagnosis and epidemiology. Trop Parasitol 2013; 3: 26-34.
110. Ramirez JD, Sanchez LV, Bautista DC, Corredor AF, Florez AC, Stensvold CR. *Blastocystis* subtypes detected in humans and animals from Colombia. Infect Genet Evol 2014; 22: 223-8.
111. Güreser AS, Comba A, Karasartova D, *et al.* Detection of *Blastocystis* subtypes in children with functional abdominal pain and celiac disease in Çorum, Turkey Iran J Parasitol 2022; 17(3): 296-305.
112. Salehi R, Rostami A, Mirjalali H, Stensvold CR, Haghighi A. Genetic characterization of *Blastocystis* from poultry, livestock animals and humans in the southwest region of Iran-Zoonotic implications. Transbound Emerg Dis 2022; 69(3): 1178-85.
113. AbuOdeh R, Ezzedine S, Samie A, Stensvold CR, ElBakri A. Prevalence and subtype distribution of *Blastocystis* in healthy individuals in Sharjah, United Arab Emirates. Infect Genet Evol 2016; 37: 158-62.
114. Alfellani MA, Jacob AS, Perea NO, *et al.* Diversity and distribution of *Blastocystis* sp. subtypes in non-human primates. Parasitology 2013: 140(8): 966-71.
115. Scanlan PD, Stensvold CR. *Blastocystis:* getting to grips with our guileful guest. Trends Parasitol 2013; 29(11): 523-9.
116. Wylezich C, Belka A, Hanke D, Beer M, Blome S, Höper D. Metagenomics for broad and improved parasite detection: a proof-of-concept study using swine faecal samples. Int J Parasitol 2019; 49(10): 769-77.
117. Stensvold CR, Lewis HC, Hammerum AM, *et al. Blastocystis:* unravelling potential risk factors and clinical significance of a common but neglected parasite. Epidemiol Infect 2009; 137(11): 1655-63.
118. Galán-Puchades MT, Trelis M, Sáez-Durán S, *et al.* One health approach to zoonotic parasites: molecular detection of intestinal protozoans in an urban population of Norway rats, *Rattus norvegicus*, in Barcelona, Spain. Pathogens 2021; 10(3): 311.
119. Katsumata M, Yoshikawa H, Tokoro M, *et al.* Molecular phylogeny of *Blastocystis* isolates from wild rodents captured in Indonesia and Japan. Parasitol Res 2018; 117(9): 2841-6.
120. Barati M, KarimiPourSaryazdi A, Rahmanian V, *et al.* Global prevalence and subtype distribution of *Blastocystis* sp. in rodents, birds, and water supplies: A systematic review and meta-analysis. Prev Vet Med 2022; 208: 105770.
121. AbuOdeh R, Ezzedine S, Madkour M, *et al.* Molecular subtyping of *Blastocystis* from diverse animals in the United Arab Emirates. Protist 2019; 170(5): 125679.
122. Liu X, Ni F, Wang R, *et al.* Occurrence and subtyping of *Blastocystis* in coypus (*Myocastor coypus*) in China. Parasit Vectors 2022; 15(1): 14.
123. Heitlinger E, Ferreira SCM, Thierer D, Hofer H, East ML. The intestinal eukaryotic and bacterial biome of spotted hyenas: the impact of social status and age on diversity and composition. Front Cell Infect Microbiol 2017; 7: 262.
124. Malheiros AF, Stensvold CR, Clark CG, Braga GB, Shaw JJ. Short report: molecular characterization of *Blastocystis* obtained from members of the indigenous Tapirapé ethnic group from the Brazilian Amazon region, Brazil. Am J Trop Med Hyg 2011; 85(6): 1050-3.
125. Salehi R, Haghighi A, Stensvold CR, *et al.* Prevalence and subtype identification of *Blastocystis* isolated from humans in Ahvaz, Southwestern Iran. Gastroenterol Hepatol Bed Bench 2017; 10(3): 235-41.
126. Mirjalali H, Abbasi MR, Naderi N, *et al.* Distribution and phylogenetic analysis of *Blastocystis* sp. subtypes isolated from IBD patients and healthy individuals in Iran. Eur J Clin Microbiol Infect Dis 2017; 36(12): 2335-42.
127. Alinaghizade A, Mirjalali H, Mohebali M, Stensvold CR, Rezaeian M. Inter- and intra-subtype variation of *Blastocystis* subtypes isolated from diarrheic and non-diarrheic patients in Iran. Infect Genet Evol 2017; 50: 77-82.

128. Pandey PK, Verma P, Marathe N, *et al.* Prevalence and subtype analysis of *Blastocystis* in healthy Indian individuals. Infect Genet Evol 2015; 31: 296-9.
129. Ozyurt M, Kurt O, Mølbak K, Nielsen HV, Haznedaroglu T, Stensvold CR. Molecular epidemiology of *Blastocystis* infections in Turkey. Parasitol Int 2008; 57(3): 300-6.
130. Jiménez PA, Jaimes JE, Ramírez JD. A summary of *Blastocystis* subtypes in North and South America. Parasit Vectors 2019; 12(1): 376.
131. Betts EL, Gentekaki E, Thomasz A, Breakell V, Carpenter AI, Tsaousis AD. Genetic diversity of *Blastocystis* in non-primate animals. Parasitology 2018; 145(9): 1228-34.
132. Šloufová M, Lhotská Z, Jirků M, *et al.* Comparison of molecular diagnostic approaches for the detection and differentiation of the intestinal protist *Blastocystis* sp. in humans. Parasite 2022; 29: 30.
133. Scanlan PD, Stensvold CR, Cotter PD. Development and application of a *Blastocystis* subtype-specific PCR assay reveals that mixed-subtype infections are common in a healthy human population. Appl Environ Microbiol 2015; 81(12): 4071-6.
134. Jepps MW, Dobell C. *Dientamoeba fragilis* n. g., n. sp., new intestinal amoeba from man. Parasitology 1918; 10: 352-67.
135. Windsor JJ, Johnson EH. *Dientamoeba fragilis:* the unflagellated human flagellate. Br J Biomed Sci 1999; 56(4): 293-306.
136. Munasinghe VS, Vella NG, Ellis JT, Windsor PA, Stark D. Cyst formation and faecal-oral transmission of *Dientamoeba fragilis* – the missing link in the life cycle of an emerging pathogen. Int J Parasitol 2013; 43(11): 879-83.
137. Lee D. The structure and development of *Histomonas meleagridis* (Mastigamoebidae: Protozoa) in the female reproductive tract of its intermediate host, *Heterakis gallinarum* (Nematoda). Parasitology 1969; 59: 877-84.
138. Röser D, Nejsum P, Carlsgart AJ, Nielsen HV, Stensvold CR. DNA of *Dientamoeba fragilis* detected within surface-sterilized eggs of *Enterobius vermicularis.* Exp Parasitol 2013; 133(1): 57-61.
139. Röser D, Simonsen J, Nielsen HV, Stensvold CR, Mølbak K. History of antimicrobial use and the risk of *Dientamoeba fragilis* infection. Eur J Clin Microbiol Infect Dis 2015; 34(6): 1145-51.
140. Stensvold CR, Arendrup MC, Mølbak K, Nielsen HV. The prevalence of *Dientamoeba fragilis* in patients with suspected enteroparasitic disease in a metropolitan area in Denmark. Clin Microbiol Infect 2007; 13(8): 839-42.
141. Stark DJ, Beebe N, Marriott D, Ellis JT, Harkness J. Dientamoebiasis: clinical importance and recent advances. Trends Parasitol 2006; 22(2): 92-6.
142. Norberg A, Nord CE, Evengård B. *Dientamoeba fragilis* – a protozoal infection which may cause severe bowel distress. Clin Microbiol Infect 2003; 9(1): 65-8.
143. Verweij JJ, Mulder B, Poell B, van Middelkoop D, Brienen EA, van Lieshout L. Real-time PCR for the detection of *Dientamoeba fragilis* in fecal samples. Mol Cell Probes 2007; 21(5-6): 400-4.
144. Röser D, Simonsen J, Nielsen HV, Stensvold CR, Mølbak K. *Dientamoeba fragilis* in Denmark: epidemiological experience derived from four years of routine real-time PCR. Eur J Clin Microbiol Infect Dis 2013; 32(10): 1303-10.
145. Roser D, Simonsen J, Stensvold CR, *et al.* Metronidazole therapy for treating dientamoebiasis in children is not associated with better clinical outcomes: a randomized, double-blinded and placebo-controlled clinical trial. Clin Infect Dis 2014; 58(12): 1692-9.
146. Jokelainen P, Hebbelstrup Jensen B, Andreassen BU, *et al. Dientamoeba fragilis* – a commensal in children in Danish day care centers. J Clin Microbiol 2017; 55(6): 1707-13.
147. Hebbelstrup Jensen B, Stensvold CR, Struve C, *et al.* Enteroaggregative *Escherichia coli* in daycare – a 1-year dynamic cohort study. Front Cell Infect Microbiol 2016; 6: 75.
148. Engsbro AL, Stensvold CR, Nielsen HV, Bytzer P. Treatment of *Dientamoeba fragilis* in patients with irritable bowel syndrome. Am J Trop Med Hyg 2012; 87(6): 1046-52.

149. Halkjaer S, Stensvold CR, Petersen AM. [*Dientamoeba fragilis* infection as cause of severe abdominal discomfort and flatulence]. Ugeskr Laeger 2015; 177(2A): 30-1.
150. Gray TJ, Kwan Y, Phan T, Robertson G, Cheong E, Gottlieb T. *Dientamoeba fragilis* – a family cluster of disease associated with marked peripheral eosinophilia. Clin Infect Dis 2013; 57(6): 845-8.
151. Lai KK, Goyne HE, Hernandez-Gonzalo D, *et al. Cystoisospora belli* infection of the gallbladder in immunocompetent patients: a clinicopathologic review of 18 cases. Am J Surg Pathol 2016; 40(8): 1070-4.
152. Röser D, Stensvold CR. Anisakiasis mistaken for dientamoebiasis? Clin Infect Dis 2013; 57(10): 1500.
153. Stensvold CR, Jespersgaard C, Dinesen K, Sorensen JF, Molbak K, Nielsen HV. Prevalence of *Dientamoeba fragilis* genotypes in different Danish populations as assessed by PCR and SNP genotyping. Trop Med Int Health 2007; 12(Suppl): 182
154. Jirků M, Kašparová A, Lhotská Z, *et al.* A cross-sectional study on the occurrence of the intestinal protist, *Dientamoeba fragilis*, in the gut-healthy volunteers and their animals. Int J Mol Sci 2022; 23(23): 15407.
155. Unalan-Altintop T, Vahabov C, Ergunay K, *et al.* Investigation of *Dientamoeba fragilis* and *Blastocystis* in patients from Turkey with ulcerative colitis and irritable bowel syndrome: Any relation with genotypes? Acta Trop 2022; 231: 106451.
156. Oliveira-Arbex AP, David É, Cacciò SM, *et al.* Prevalence and genetic characterization of *Dientamoeba fragilis* in asymptomatic children attending daycare centers. Rev Inst Med Trop Sao Paulo 2021; 63: e39.
157. Cacciò SM, Sannella AR, Bruno A, *et al.* Multilocus sequence typing of *Dientamoeba fragilis* identified a major clone with widespread geographical distribution. Int J Parasitol 2016; 46(12): 793-8.
158. Johnson JA, Clark CG. Cryptic genetic diversity in *Dientamoeba fragilis*. J Clin Microbiol 2000; 38(12): 4653-4.
159. Peek R, Reedeker FR, van Gool T. Direct amplification and genotyping of *Dientamoeba fragilis* from human stool specimens. J Clin Microbiol 2004; 42(2): 631-5.
160. Stark D, Beebe N, Marriott D, Ellis J, Harkness J. Prospective study of the prevalence, genotyping, and clinical relevance of *Dientamoeba fragilis* infections in an Australian population. J Clin Microbiol 2005; 43(6): 2718-23.
161. Windsor JJ, Macfarlane L, Clark CG. Internal transcribed spacer dimorphism and diversity in *Dientamoeba fragilis.* J Eukaryot Microbiol 2006; 53(3): 188-92.
162. Bart A, van der Heijden HM, Greve S, Speijer D, Landman WJ, van Gool T. Intragenomic variation in the internal transcribed spacer 1 region of *Dientamoeba fragilis* as a molecular epidemiological marker. J Clin Microbiol 2008; 46(10): 3270-5.
163. Stark D, Barratt J, Ellis J, Harkness J, Marriott D. Repeated *Dientamoeba fragilis* infections: a case report of two families from Sydney, Australia. Infect Dis Rep 2009; 1(1): e4.
164. David É, Guimarães S, de Oliveira AP, *et al.* Molecular characterization of intestinal protozoa in two poor communities in the State of São Paulo, Brazil. Parasit Vectors 2015; 8: 103.
165. Caccio SM. Molecular epidemiology of *Dientamoeba fragilis.* Acta Trop 2018; 184: 73-7.
166. Bilic I, Jaskulska B, Souillard R, Liebhart D, Hess M. Multi-locus typing of *Histomonas meleagridis* isolates demonstrates the existence of two different genotypes. PLoS One 2014; 9(3): e92438.
167. Ogunniyi T, Balogun H, Shasanya B. Ectoparasites and endoparasites of peridomestic house-rats in ile-ife, Nigeria and implication on human health. Iran J Parasitol 2014; 9(1): 134-40.
168. Helenbrook WD, Wade SE, Shields WM, Stehman SV, Whipps CM. Gastrointestinal parasites of Ecuadorian mantled howler monkeys (*Alouatta palliata aequatorialis*) based on fecal analysis. J Parasitol 2015; 101(3): 341-50.
169. Myers BJ, Kuntz RE. Intestinal protozoa of the baboon *Papio doguera* Pucheran, 1856. J Protozool 1968; 15(2): 363-5.

170. Hegner R, Chu HJ. A comparative study of the intestinal protozoa of wild monkeys and man. Am J Epidemiol 1930; 12(1): 62-108.
171. Sing A, Tybus K, Heesemann J, Mathis A. Molecular diagnosis of an *Enterocytozoon bieneusi* human genotype C infection in a moderately immunosuppressed human immunodeficiency virus seronegative liver-transplant recipient with severe chronic diarrhea. J Clin Microbiol 2001; 39(6): 2371-2.
172. ten Hove RJ, Van Lieshout L, Beadsworth MB, *et al.* Characterization of genotypes of *Enterocytozoon bieneusi* in immunosuppressed and immunocompetent patient groups. J Eukaryot Microbiol 2009; 56(4): 388-93.
173. Sak B, Petrzelkova KJ, Kvetonova D, *et al.* Diversity of microsporidia, *Cryptosporidium* and *Giardia* in mountain gorillas (*Gorilla beringei beringei*) in Volcanoes National Park, Rwanda. PLoS One 2014; 9(11): e109751.
174. Sak B, Kváč M, Květoňová D, Albrecht T, Piálek J. The first report on natural *Enterocytozoon bieneusi* and *Encephalitozoon* spp. infections in wild east-European house mice (*Mus musculus musculus*) and west-European house mice (*M. m. domesticus*) in a hybrid zone across the Czech Republic-Germany border. Vet Parasitol 2011; 178(3-4): 246-50.
175. Vioque F, Dashti A, Santin M, *et al.* Wild micromammal host spectrum of zoonotic eukaryotic parasites in Spain. Occurrence and genetic characterisation. Transbound Emerg Dis 2022; 69(5): e2926-42.
176. Ptackova E, Kostygov AY, Chistyakova LV, *et al.* Evolution of Archamoebae: morphological and molecular evidence for pelobionts including *Rhizomastix*, *Entamoeba*, *Iodamoeba*, and *Endolimax*. Protist 2013; 164(3): 380-410.
177. León-Avila G, Tovar J. Mitosomes of *Entamoeba histolytica* are abundant mitochondrion-related remnant organelles that lack a detectable organellar genome. Microbiology (Reading) 2004; 150(Pt 5): 1245-50.
178. Levecke B, Dorny P, Vercammen F, *et al.* Transmission of *Entamoeba nuttalli* and *Trichuris trichiura* from nonhuman primates to humans. Emerg Infect Dis 2015; 21(10): 1871-2.
179. Royer TL, Gilchrist C, Kabir M, *et al. Entamoeba bangladeshi* nov. sp., Bangladesh. Emerg Infect Dis 2012; 18(9): 1543-5.
180. Ngobeni R, Samie A, Moonah S, Watanabe K, Petri WA, Gilchrist C. *Entamoeba* species in South Africa: correlations with the host microbiome, parasite burdens, and first description of *Entamoeba bangladeshi* outside of Asia. J Infect Dis 2017; 216(12): 1592-600.
181. Clark CG, Stensvold CR. The continuously expanding universe of *Entamoeba*. In: Nozaki T, Bhattacharya A. Amebiasis. Tokyo: Springer, 2015:9-25.
182. Vlčková K, Kreisinger J, Pafčo B, et al. Diversity of *Entamoeba* spp. in African great apes and humans: an insight from Illumina MiSeq high-throughput sequencing. Int J Parasitol 2018; 48(7): 519-30.
183. Zahedi A, Greay TL, Paparini A, Linge KL, Joll CA, Ryan UM. Identification of eukaryotic microorganisms with 18S rRNA next-generation sequencing in wastewater treatment plants, with a more targeted NGS approach required for *Cryptosporidium* detection. Water Res 2019; 158: 301-12.
184. Moreno Y, Moreno-Mesonero L, Amorós I, Pérez R, Morillo JA, Alonso JL. Multiple identification of most important waterborne protozoa in surface water used for irrigation purposes by 18S rRNA amplicon-based metagenomics. Int J Hyg Environ Health 2018; 221(1): 102-11.
185. Clark CG, Kaffashian F, Tawari B, *et al.* New insights into the phylogeny of *Entamoeba* species provided by analysis of four new small-subunit rRNA genes. Int J Syst Evol Microbiol 2006; 56(Pt 9): 2235-9.
186. Sangaré I, Guiguemdé KT, Zida A, *et al.* Prevalence of intestinal parasitic infections among pregnant women in Bobo-Dioulasso (Burkina Faso). Ann Parasitol 2021; 67(3): 489-97.
187. Araujo Z, Brandes S, Pinelli E, *et al.* Seropositivity for ascariosis and toxocariosis and cytokine expression among the indigenous people in the Venezuelan Delta region. Rev Inst Med Trop Sao Paulo 2015; 57(1): 47-55.

188. Incani RN, Ferrer E, Hoek D, *et al.* Diagnosis of intestinal parasites in a rural community of Venezuela: advantages and disadvantages of using microscopy or RT-PCR. Acta Trop 2017; 167: 64-70.
189. Eiset AH, Stensvold CR, Fuursted K, Nielsen HV, Wejse C. High prevalence of methicillin-resistant *Staphylococcus aureus*, *Giardia*, and *Blastocystis* in asymptomatic Syrian asylum seekers in Denmark during 2016 through 2018. J Migr Health 2020; 1-2: 100016.
190. Engsbro AL, Stensvold CR, Nielsen HV, Bytzer P. Prevalence, incidence, and risk factors of intestinal parasites in Danish primary care patients with irritable bowel syndrome. Scand J Infect Dis 2014; 46(3): 204-9.
191. Dobell C, Jepps MW. On the three common intestinal Entamoebae of man, and their differential diagnosis. Br Med J 1917; 1(2941): 607-12.
192. Chihi A, Stensvold CR, Ben-Abda I, *et al.* Development and evaluation of molecular tools for detecting and differentiating intestinal amoebae in healthy individuals. Parasitology 2019; 146(6): 821-7.
193. Matey EJ, Tokoro M, Nagamoto T, *et al.* Lower prevalence of *Entamoeba* species in children with vertically transmitted HIV infection in Western Kenya. AIDS 2016; 30(5): 803-5.
194. Matsumura T, Hendarto J, Mizuno T, *et al.* Possible pathogenicity of commensal *Entamoeba hartmanni* revealed by molecular screening of healthy school children in Indonesia. Trop Med Health 2019; 47: 7.
195. Dobell C. The amoebae living in man. A Zoologicla Monograph. London, England.: John Bale, Sons & Danielson, 1919.
196. Matthews JR. A mensurative study of the cysts of *Entamoeba coli*. Ann Trop Med Parasitol 1919; 12: 259-72.
197. Nolan MJ, Unger M, Yeap YT, *et al.* Molecular characterisation of protist parasites in human-habituated mountain gorillas (*Gorilla beringei beringei*), humans and livestock, from Bwindi impenetrable National Park, Uganda. Parasit Vectors 2017; 10(1): 340.
198. Calegar DA, Monteiro KJL, Bacelar PAA, *et al.* Epidemiology, species composition and genetic diversity of tetra- and octonucleated *Entamoeba* spp. in different Brazilian biomes. Parasit Vectors 2021; 14(1): 160.
199. Kessel JF. Intestinal protozoa of the domestic pig. Am J Trop Med Hyg 1928; 8: 481-501.
200. Pinilla JC, Morales E, Muñoz AAF. A survey for potentially zoonotic parasites in backyard pigs in the Bucaramanga metropolitan area, northeast Colombia. Vet World 2021; 14(2): 372-9.
201. Mendoza-Gómez M, Pulido-Villamarín A, Barbosa-Buitrago A, Aranda-Silva M. Presence of gastrointestinal parasites in swine and human of four swine production farms in Cundinamarca-Colombia. Revista MVZ Córdoba 2015; 20(1): 5014-27.
202. Dibakou SE, Ngoubangoye B, Boundenga L, *et al.* Preliminary assessment of gastrointestinal parasites of the sun-tailed monkey (*Allochrocebus solatus*) in a semi-free-ranging colony. J Med Primatol 2022; 51(3): 127-33.
203. Ling Z, Feng M, Xu J, Tachibana H, Cheng X. Identification of Mamu-DRB1 gene as a susceptibility factor for *Entamoeba nuttalli* infection in Chinese *Macaca mulatta*. Infect Genet Evol 2021; 93: 104952.
204. Kumar S, Kumara HN, Santhosh K, Sundararaj P. Prevalence of gastrointestinal parasites in lion-tailed macaque *Macaca silenus* in central Western Ghats, India. Primates 2019; 60(6): 537-46.
205. Howells ME, Pruetz J, Gillespie TR. Patterns of gastro-intestinal parasites and commensals as an index of population and ecosystem health: the case of sympatric western chimpanzees (*Pan troglodytes verus*) and guinea baboons (*Papio hamadryas papio*) at Fongoli, Senegal. Am J Primatol 2011; 73(2): 173-9.
206. Dibakou SE, Maloueki U, Ngoubangoye B, *et al.* Diversity of gastrointestinal parasites in sympatric mammals in Moukalaba-Doudou National Park, Gabon. Vet World 2021; 14(12): 3149-55.

207. Sleeman JM, Meader LL, Mudakikwa AB, Foster JW, Patton S. Gastrointestinal parasites of mountain gorillas (*Gorilla gorilla beringei*) in the Parc National des Volcans, Rwanda. J Zoo Wildl Med 2000; 31(3): 322-8.
208. Jirků-Pomajbíková K, Čepička I, Kalousová B, *et al.* Molecular identification of *Entamoeba* species in savanna woodland chimpanzees (*Pan troglodytes schweinfurthii*). Parasitology 2016: 1-8.
209. Kessel JF. Experimental infection of rats and mice with the common intestinal amoebae of man. University of California Publications in Zoology 1923; 20: 409-30.
210. Regendanz P. Über die Übertragung der *Entamoeba histolytica*, *Entamoeba coli* und *Dientamoeba fragilis* auf Ratten. Zentralbl Bakteriol Parasitenk Infektionskrank Hyg 1929; 111: 412-9.
211. Neal RA. An experimental study of *Entamoeba muris* (Grassi, 1879); its morphology, affinities and host-parasite relationship. Parasitology 1950; 40(3-4): 343-65.
212. Levine ND. Protozoan parasites of domestic animals and man. Minnesota: Burgess Publishing Company, 1961.
213. Sargeaunt PG, Williams JE. Electrophoretic isoenzyme patterns of *Entamoeba histolytica* and *Entamoeba coli*. Trans R Soc Trop Med Hyg 1978; 72(2): 164-6.
214. Sargeaunt PG, Williams JE, Grene JD. The differentiation of invasive and non-invasive *Entamoeba histolytica* by isoenzyme electrophoresis. Trans R Soc Trop Med Hyg 1978; 72(5): 519-21.
215. Tannich E, Horstmann RD, Knobloch J, Arnold HH. Genomic DNA differences between pathogenic and nonpathogenic *Entamoeba histolytica*. Proc Natl Acad Sci U S A 1989; 86(13): 5118-22.
216. Tannich E, Bruchhaus I, Walter RD, Horstmann RD. Pathogenic and nonpathogenic *Entamoeba histolytica:* identification and molecular cloning of an iron-containing superoxide dismutase. Mol Biochem Parasitol 1991; 49(1): 61-71.
217. Verweij JJ, Blangé RA, Templeton K, *et al.* Simultaneous detection of *Entamoeba histolytica*, *Giardia lamblia*, and *Cryptosporidium parvum* in fecal samples by using multiplex real-time PCR. J Clin Microbiol 2004; 42(3): 1220-3.
218. Efunshile MA, Ngwu BA, Kurtzhals JA, Sahar S, König B, Stensvold CR. Molecular detection of the carriage rate of four intestinal protozoa with real-time polymerase chain reaction: possible overdiagnosis of *Entamoeba histolytica* in Nigeria. Am J Trop Med Hyg 2015; 93(2): 257-62.
219. Stensvold CR, Nielsen HV. Comparison of microscopy and PCR for the detection of intestinal parasites in Danish patients supports incentive for molecular screening platforms. J Clin Microbiol 2012; 50(2): 540-1.
220. Abozahra R, Mokhles M, Baraka K. Prevalence and molecular differentiation of *Entamoeba histolytica*, *Entamoeba dispar*, *Entamoeba moshkovskii*, and *Entamoeba hartmanni* in Egypt. Acta Parasitol 2020; 65(4): 929-35.
221. Samie A, Mahlaule L, Mbati P, Nozaki T, ElBakri A. Prevalence and distribution of *Entamoeba* species in a rural community in northern South Africa. Food Waterborne Parasitol 2020; 18: e00076.
222. da Silva CAV, de Oliveira IMC, Cruz RE, *et al.* South American *Entamoeba dispar* strains produce amoebic liver abscesses with different pathogenicities and evolutionary kinetics. Acta Trop 2021; 224: 106114.
223. Guzmán-Silva MA, Santos HL, Peralta RS, Peralta JM, de Macedo HW. Experimental amoebic liver abscess in hamsters caused by trophozoites of a Brazilian strain of *Entamoeba dispar*. Exp Parasitol 2013; 134(1): 39-47.
224. Dolabella SS, Serrano-Luna J, Navarro-García F, *et al.* Amoebic liver abscess production by *Entamoeba dispar*. Ann Hepatol 2012; 11(1): 107-17.
225. Köster PC, Lapuente J, Pizarro A, *et al.* Presence and genetic diversity of enteric protists in captive and semi-captive non-human primates in côte d'Ivoire, Sierra Leone, and Peru. Int J Parasitol Parasites Wildl 2022; 17: 26-34.

226. Dong H, Li J, Qi M, *et al.* Prevalence, molecular epidemiology, and zoonotic potential of *Entamoeba* spp. in nonhuman primates in China. Infect Genet Evol 2017; 54: 216-20.
227. Zaffino M, Dubar M, Debourgogne A, Bisson C, Machouart M. Development of a new TaqMan PCR assay for the detection of both *Entamoeba gingivalis* genotypes. Diagn Microbiol Infect Dis 2019; 95(4): 114886.
228. Sharifi M, Jahanimoghadam F, Babaei Z, *et al.* Prevalence and associated-factors for *Entamoeba gingivalis* in adolescents in southeastern Iran by culture and PCR, 2017. Iran J Public Health 2020; 49(2): 351-9.
229. Noble ER, Noble GA. Parasitology – The biology of animal parasites. 4. ed. Philadelphia: Lea & Febiger, 1976.
230. Azadbakht K, Baharvand P, Artemes P, Niazi M, Mahmoudvand H. Prevalence and risk factors of oral cavity parasites in pregnant women in Western Iran. Parasite Epidemiol Control 2022; 19: e00275.
231. Yaseen A, Mahafzah A, Dababseh D, *et al.* Oral colonization by *Entamoeba gingivalis* and *Trichomonas tenax:* a PCR-based study in health, gingivitis, and periodontitis. Front Cell Infect Microbiol 2021; 11: 782805.
232. Keeler EL, Merenstein C, Reddy S, *et al.* Widespread, human-associated redondoviruses infect the commensal protozoan *Entamoeba gingivalis*. Cell Host Microbe 2023; 31(1): 58-68.e5
233. Koller G, Foschi F, Mitchell P, Witherden E, Bruce K, Mannocci F. Amoebae in chronic, polymicrobial endodontic infections are associated with altered microbial communities of increased virulence. J Clin Med 2020; 9(11): 3700.
234. Garcia G, Ramos F, Martinez-Hernandez F, Hernandez L, Yanez J, Gaytan P. A new subtype of *Entamoeba gingivalis:* "*E. gingivalis* ST2, kamaktli variant". Parasitol Res 2018; 117(4): 1277-84.
235. Yamamoto A, Kikuta N, Hashimoto T, Oyaizu H, Goto N. Nucleotide sequence of the SrRNA gene of *Entamoeba gingivalis:* applications for construction of a species-specific DNA probe and phylogenetic analysis. Microbiol Immunol 1995; 39(3): 185-92.
236. Martin-Garcia DF, Sallam M, Garcia G, Santi-Rocca J. Parasites in periodontal health and disease: a systematic review and meta-analysis. Adv Exp Med Biol 2022; 1373: 95-111.
237. Randolph NK, McAloney CA, Ossiboff R, Hernandez B, Cook M, Hokamp JA. Nasal colonization by *Entamoeba gingivalis* in a 13-year-old Italian greyhound. Vet Clin Pathol 2022; 51(2): 269-72.
238. Rousset JJ, Laissus B, Michel C. [Preliminary data on canine enzootics due to *Trichomonas canistomae* and *Entamoeba gingivalis* in Parisian dogs]. C R Seances Soc Biol Fil 1970; 164(12): 2468-9.
239. Walter Beck J, Barrett-Connor. Medical parasitology: Mosby, 1971.
240. Blacklock D, Southwell T. A guide to human parasitology – for medical practitioners. London: H. K. Lewis & Co. Ltd., 1977.
241. Spencer FM, Monroe LS. The color atlas of intestinal parasites. Illinois: Charles C Thomas, 1961.
242. Chang AM, Chen CC, Huffman MA. *Entamoeba* spp. in wild formosan rock macaques (*Macaca cyclopis*) in an area with frequent human-macaque contact. J Wildl Dis 2019; 55(3): 608-18.
243. Hamad I, Sokhna C, Raoult D, Bittar F. Molecular detection of eukaryotes in a single human stool sample from Senegal. PLoS One 2012; 7(7): e40888.
244. Ali IK, Hossain MB, Roy S, *et al. Entamoeba moshkovskii* infections in children, Bangladesh. Emerg Infect Dis 2003; 9(5): 580-4.
245. Fotedar R, Stark D, Beebe N, Marriott D, Ellis J, Harkness J. PCR detection of *Entamoeba histolytica*, *Entamoeba dispar*, and *Entamoeba moshkovskii* in stool samples from Sydney, Australia. J Clin Microbiol 2007; 45(3): 1035-7.
246. López MC, León CM, Fonseca J, *et al.* Molecular epidemiology of *Entamoeba*: first description of *Entamoeba moshkovskii* in a rural area from central Colombia. PLoS One 2015; 10(10): e0140302.

247. Sardar SK, Das K, Maruf M, *et al.* Molecular evidence suggests the occurrence of *Entamoeba moshkovskii* in pigs with zoonotic potential from eastern India. Folia Parasitol (Praha) 2022; 69: 2022.012.
248. Parija SC, Khairnar K. *Entamoeba moshkovskii* and *Entamoeba dispar*-associated infections in Pondicherry, India. J Health Popul Nutr 2005; 23(3): 292-5.
249. Shimokawa C, Kabir M, Taniuchi M, *et al. Entamoeba moshkovskii* is associated with diarrhea in infants and causes diarrhea and colitis in mice. J Infect Dis 2012; 206(5): 744-51.
250. Singh A, Banerjee T, Khan U, Shukla SK. Epidemiology of clinically relevant *Entamoeba* spp. (*E. histolytica*/*dispar*/*moshkovskii*/*bangladeshi*): a cross sectional study from North India. PLoS Negl Trop Dis 2021; 15(9): e0009762.
251. Ali IKM, Roy S. A real-time PCR assay for simultaneous detection and differentiation of four common *Entamoeba* species that infect humans J Clin Microbiol 2020; 59(1): e01986-20.
252. Sardar SK, Kobayashi S, Das K, *et al.* Development of a simple PCR-RFLP technique for detection and differentiation of *E. histolytica*, *E. dispar* and *E. moshkovskii*. Parasitol Res 2023; 122(1): 139-44.
253. Ngui R, Angal L, Fakhrurrazi SA, *et al.* Differentiating *Entamoeba histolytica*, *Entamoeba dispar* and *Entamoeba moshkovskii* using nested polymerase chain reaction (PCR) in rural communities in Malaysia. Parasit Vectors 2012; 5: 187.
254. Burrows RB, Klink GE. *Endamoeba polecki* infections in man. Am J Hyg 1955; 62(2): 156-67.
255. Desowitz RS, Barnish G. *Entamoeba polecki* and other intestinal protozoa in Papua New Guinea Highland children. Ann Trop Med Parasitol 1986; 80(4): 399-402.
256. Park SK, Kim DH, Deung YK, *et al.* Status of intestinal parasite infections among children in Bat Dambang, Cambodia. Korean J Parasitol 2004; 42(4): 201-3.
257. Verweij JJ, Polderman AM, Clark CG. Genetic variation among human isolates of uninucleated cyst-producing *Entamoeba* Species. J Clin Microbiol 2001; 39(4): 1644-6.
258. Pánek T, Zadrobílková E, Walker G, *et al.* First multigene analysis of Archamoebae (Amoebozoa: Conosa) robustly reveals its phylogeny and shows that Entamoebidae represents a deep lineage of the group. Mol Phylogenet Evol 2016; 98: 41-51.
259. Kuenen W, Swellengrebel N. Korte beschrijving van enkelc minder bekendc protozo uit denmenschelijkcn darm. Tijdschr v Nederl 1917: 496.
260. Wenyon C, O'Conner F. Human intestinal protozoa in the Near East. London: Bale, Sons and Danielsson, Ltd., 1917.
261. de la Luz Galvan-Ramirez M, Madriz-Elisondo AL, Ramirez CGT, de Jesus Romero Rameno J, de la OCDA, Lopez MAC. Enteroparasitism and risk factors associated with clinical manifestations in children and adults of Jalisco State in Western Mexico. Osong Public Health Res Perspect 2019; 10(1): 39-48.
262. Ramírez JD, Flórez C, Olivera M, Bernal MC, Giraldo JC. *Blastocystis* subtyping and its association with intestinal parasites in children from different geographical regions of Colombia. PLoS One 2017; 12(2): e0172586.
263. Faria CP, Zanini GM, Dias GS, *et al.* Geospatial distribution of intestinal parasitic infections in Rio de Janeiro (Brazil) and its association with social determinants. PLoS Negl Trop Dis 2017; 11(3): e0005445.
264. Raso G, Utzinger J, Silué KD, *et al.* Disparities in parasitic infections, perceived ill health and access to health care among poorer and less poor schoolchildren of rural Côte d'Ivoire. Trop Med Int Health 2005; 10(1): 42-57.
265. Constenla M, Padrós F, Palenzuela O. *Endolimax piscium* sp. nov. (Amoebozoa), causative agent of systemic granulomatous disease of cultured sole, *Solea senegalensis* Kaup. J Fish Dis 2014; 37(3): 229-40.
266. Yoshida N, Kobayashi S, Suzuki J, *et al.* Growth-promoting effects of the hydrogen-sulfide compounds produced by *Desulfovibrio desulfuricans* subsp. *desulfuricans* co-cultured with *Escherichia coli* (DH5α) on the growth of *Entamoeba* and *Endolimax* species isolates from swine. Biosci Trends 2019; 13(5): 402-10.

267. Mann AE, Mazel F, Lemay MA, *et al.* Biodiversity of protists and nematodes in the wild nonhuman primate gut. ISME J 2020; 14(2): 609-22.
268. Zanzani SA, Gazzonis AL, Epis S, Manfredi MT. Study of the gastrointestinal parasitic fauna of captive non-human primates (*Macaca fascicularis*). Parasitol Res 2016; 115(1): 307-12.
269. Kouassi RY, McGraw SW, Yao PK, *et al.* Diversity and prevalence of gastrointestinal parasites in seven non-human primates of the Taï National Park, Côte d'Ivoire. Parasite 2015; 22: 1.
270. Cancrini G, Bartoloni A, Nunez L, Paradisi F. Intestinal parasites in the Camiri, Gutierrez and Boyuibe areas, Santa Cruz Department, Bolivia. Parassitologia 1988; 30(2-3): 263-9.
271. Pena-Quistial MG, Benavides-Montano JA, Duque NJR, Benavides-Montano GA. Prevalence and associated risk factors of Intestinal parasites in rural high-mountain communities of the Valle del Cauca-Colombia. PLoS Negl Trop Dis 2020; 14(10): e0008734.
272. Maco Flores V, Marcos Raymundo LA, Terashima Iwashita A, Samalvides Cuba F, Gotuzzo Herencia E. [Distribution of entero-parasitic infections in the Peruvian Highland: study carried out in six rural communities of the department of Puno, Peru]. Rev Gastroenterol Peru 2002; 22(4): 304-9.
273. Cabrera M, Verástegui M, Cabrera R. [Prevalence of entero-parasitosis in one Andean community in the Province of Victor Fajardo, Ayacucho, Peru]. Rev Gastroenterol Peru 2005; 25(2): 150-5.
274. Hamad I, Keita MB, Peeters M, Delaporte E, Raoult D, Bittar F. Pathogenic eukaryotes in gut microbiota of western lowland gorillas as revealed by molecular survey. Sci Rep 2014; 4: 6417.
275. Pafco B, Tehlarova Z, Jirku Pomajbikova K, *et al.* Gastrointestinal protists and helminths of habituated agile mangabeys (*Cercocebus agilis*) at Bai Hokou, Central African Republic. Am J Primatol 2018; 80(2).
276. Ray HN, Banik DC. A hyper-parasite of *Iodamoeba buetschlii* from the rhesus monkey, *Macaca mulatta*, and the domestic pig. J Protozool 1965; 12: 70-2.
277. Solaymani-Mohammadi S, Rezaian M, Hooshyar H, Mowlavi GR, Babaei Z, Anwar MA. Intestinal protozoa in wild boars (*Sus scrofa*) in western Iran. J Wildl Dis 2004; 40(4): 801-3.
278. Pakandl M. The prevalence of intestinal protozoa in wild and domestic pigs. Vet Med (Praha) 1994; 39(7): 377-80.
279. Levecke B, Dorny P, Geurden T, Vercammen F, Vercruysse J. Gastrointestinal protozoa in non-human primates of four zoological gardens in Belgium. Vet Parasitol 2007; 148(3-4): 236-46.
280. Tachibana H, Cheng XJ, Kobayashi S, Fujita Y, Udono T. *Entamoeba dispar*, but not *E. histolytica*, detected in a colony of chimpanzees in Japan. Parasitol Res 2000; 86(7): 537-41.
281. Belkessa S, Ait-Salem E, Laatamna A, *et al.* Prevalence and clinical manifestations of *Giardia intestinalis* and other intestinal parasites in children and adults in Algeria. Am J Trop Med Hyg 2021; 104(3): 910-6.
282. Dib JR, Fernandez-Zenoff MV, Oquilla J, Lazarte S, Gonzalez SN. Prevalence of intestinal parasitic infection among children from a shanty town in Tucuman, Argentina. Trop Biomed 2015; 32(2): 210-5.
283. Jeong ES, Park JH, Ryu SH, *et al.* Detection of *Chilomastix mesnili* in common marmoset (*Callithrix jacchus*) and treatment with netronidazole. Iran J Parasitol 2019; 14(2): 334-9.
284. Kuze N, Kanamori T, Malim TP, *et al.* Parasites found from the feces of Bornean orangutans in Danum Valley, Sabah, Malaysia, with a redescription of *Pongobius hugoti* and the description of a new species of *Pongobius* (Nematoda: Oxyuridae). J Parasitol 2010; 96(5): 954-60.
285. Cepicka I, Kostka M, Uzlikova M, Kulda J, Flegr J. Non-monophyly of Retortamonadida and high genetic diversity of the genus *Chilomastix* suggested by analysis of SSU rDNA. Mol Phylogenet Evol 2008; 48(2): 770-5.
286. Lokmer A, Cian A, Froment A, *et al.* Use of shotgun metagenomics for the identification of protozoa in the gut microbiota of healthy individuals from worldwide populations with various industrialization levels. PLoS One 2019; 14(2): e0211139.
287. Kolisko M, Cepicka I, Hampl V, *et al.* Molecular phylogeny of diplomonads and enteromonads based on SSU rRNA, alpha-tubulin and HSP90 genes: implications for the evolutionary history of the double karyomastigont of diplomonads. BMC Evol Biol 2008; 8: 205.

288. Kasssem HH, Zaed HA, Sadaga GA. Intestinal parasitic infection among children and neonatus admitted to Ibn-Sina Hospital, Sirt, Libya. J Egypt Soc Parasitol 2007; 37(2): 371-80.
289. Hendarto J, Mizuno T, Hidayati APN, *et al.* Three monophyletic clusters in *Retortamonas* species isolated from vertebrates. Parasitol Int 2019; 69: 93-8.
290. Fussy Z, Vinopalova M, Treitli SC, *et al.* Retortamonads from vertebrate hosts share features of anaerobic metabolism and pre-adaptations to parasitism with diplomonads. Parasitol Int 2021; 82: 102308.
291. Li WC, Wang K, Li Y, Zhao LP, Xiao Y, Gu YF. Survey and molecular characterization of trichomonads in pigs in Anhui province, east China, 2014. Iran J Parasitol 2018; 13(4): 602-10.
292. Li WC, Wang K, Zhang W, Wu J, Gu YF, Zhang XC. Prevalence and molecular characterization of intestinal trichomonads in pet dogs in aast China. Korean J Parasitol 2016; 54(6): 703-10.
293. Li WC, Ying M, Gong PT, *et al. Pentatrichomonas hominis:* prevalence and molecular characterization in humans, dogs, and monkeys in northern China. Parasitol Res 2016; 115(2): 569-74.
294. Li X, Li J, Zhang X, Yang Z, Yang J, Gong P. Prevalence of *Pentatrichomonas hominis* infections in six farmed wildlife species in Jilin, China. Vet Parasitol 2017; 244: 160-3.
295. Kamaruddin M, Tokoro M, Rahman MM, *et al.* Molecular characterization of various trichomonad species isolated from humans and related mammals in Indonesia. Korean J Parasitol 2014; 52(5): 471-8.
296. Matthew MA, Yang N, Ketzis J, Mukaratirwa S, Yao C. *Trichomonas tenax:* a neglected protozoan infection in the oral cavities of humans and dogs – a scoping review. Trop Med Infect Dis 2023; 8(1): 60.
297. O'Brien Andersen L, Karim AB, Roager HM, *et al.* Associations between common intestinal parasites and bacteria in humans as revealed by qPCR. Eur J Clin Microbiol Infect Dis 2016; 35(9): 1427-31.
298. Petersen AM, Stensvold CR, Mirsepasi H, *et al.* Active ulcerative colitis associated with low prevalence of *Blastocystis* and *Dientamoeba fragilis* infection. Scand J Gastroenterol 2013; 48(5): 638-9.
299. Rinninella E, Raoul P, Cintoni M, *et al.* What is the healthy gut microbiota composition? A changing ecosystem across age, environment, diet, and diseases. Microorganisms 2019; 7(1): 14.
300. Aggarwal N, Kitano S, Puah GRY, Kittelmann S, Hwang IY, Chang MW. Microbiome and Human Health: Current Understanding, Engineering, and Enabling Technologies. Chem Rev 2023; 123(1): 31-72.
301. Iebba V, Santangelo F, Totino V, *et al.* Gut microbiota related to *Giardia duodenalis*, *Entamoeba* spp. and *Blastocystis hominis* infections in humans from Côte d'Ivoire. J Infect Dev Ctries 2016; 10(9): 1035-41.
302. Morton ER, Lynch J, Froment A, *et al.* Variation in rural African gut microbiota is strongly correlated with colonization by *Entamoeba* and subsistence. PLoS Genet 2015; 11(11): e1005658.
303. Gotfred-Rasmussen H, Stensvold CR, Ingham AC, *et al.* Impact of metronidazole treatment and *Dientamoeba fragilis* colonization on gut microbiota diversity. J Pediatr Gastroenterol Nutr 2021; 73(1): 23-9.
304. Stensvold CR. Thinking *Blastocystis* out of the box. Trends Parasitol 2012; 28(8): 305.
305. Rook GA, Raison CL, Lowry CA. Microbial 'old friends', immunoregulation and socioeconomic status. Clin Exp Immunol 2014; 177(1): 1-12.
306. Parfrey LW, Walters WA, Lauber CL, *et al.* Communities of microbial eukaryotes in the mammalian gut within the context of environmental eukaryotic diversity. Front Microbiol 2014; 5: 298.
307. Weinstock JV, Elliott DE. Helminths and the IBD hygiene hypothesis. Inflamm Bowel Dis 2009; 15(1): 128-33.

308. Ni J, Wu GD, Albenberg L, Tomov VT. Gut microbiota and IBD: causation or correlation? Nat Rev Gastroenterol Hepatol 2017; 14(10): 573-84.
309. Rigottier-Gois L. Dysbiosis in inflammatory bowel diseases: the oxygen hypothesis. ISME J 2013; 7(7): 1256-61.
310. Putignani L, Del Chierico F, Vernocchi P, *et al.* Gut microbiota dysbiosis as risk and premorbid factors of IBD and IBS along the childhood-adulthood transition. Inflamm Bowel Dis 2016; 22(2): 487-504.
311. Walker AW, Sanderson JD, Churcher C, *et al.* High-throughput clone library analysis of the mucosa-associated microbiota reveals dysbiosis and differences between inflamed and non-inflamed regions of the intestine in inflammatory bowel disease. BMC Microbiol 2011; 11: 7.
312. Parfrey LW, Walters WA, Knight R. Microbial eukaryotes in the human microbiome: ecology, evolution, and future directions. Front Microbiol 2011; 2: 153.
313. Kittelmann S, Devente SR, Kirk MR, Seedorf H, Dehority BA, Janssen PH. Phylogeny of intestinal ciliates, including *Charonina ventriculi*, and comparison of microscopy and 18S rRNA gene pyrosequencing for rumen ciliate community structure analysis. Appl Environ Microbiol 2015; 81(7): 2433-44.
314. Mishra P, Tulsani NJ, Jakhesara SJ, *et al.* Exploring the eukaryotic diversity in rumen of Indian camel (*Camelus dromedarius*) using 18S rRNA amplicon sequencing. Arch Microbiol 2020; 202(7): 1861-72.
315. Cammarota G, Ianiro G, Tilg H, *et al.* European consensus conference on faecal microbiota transplantation in clinical practice. Gut 2017; 66(4): 569-80.
316. Terveer EM, van Gool T, Ooijevaar RE, *et al.* Human transmission of *Blastocystis* by fecal microbiota transplantation without development of gastrointestinal symptoms in recipients. Clin Infect Dis 2020; 71(10): 2630-6.
317. Chouari R, Leonard M, Bouali M, et al. Eukaryotic molecular diversity at different steps of the wastewater treatment plant process reveals more phylogenetic novel lineages. World J Microbiol Biotechnol 2017; 33(3): 44.
318. Matsunaga K, Kubota K, Harada H. Molecular diversity of eukaryotes in municipal wastewater treatment processes as revealed by 18S rRNA gene analysis. Microbes Environ 2014; 29(4): 401-7.
319. Hartmeyer GN, Stensvold CR, Fabricius T, *et al. Plasmodium cynomolgi* as cause of malaria in tourist to southeast Asia, 2018. Emerg Infect Dis 2019; 25(10): 1936-9.
320. Huber M, Koller B, Gitler C, *et al. Entamoeba histolytica* ribosomal RNA genes are carried on palindromic circular DNA molecules. Mol Biochem Parasitol 1989; 32(2-3): 285-96.
321. Villalobos G, Orozco-Mosqueda GE, Lopez-Perez M, *et al.* Suitability of internal transcribed spacers (ITS) as markers for the population genetic structure of *Blastocystis* spp. Parasit Vectors 2014; 7: 461.
322. Moura RGF, Oliveira-Silva MB, Pedrosa AL, Nascentes GAN, Cabrine-Santos M. Occurrence of *Blastocystis* spp. in domestic animals in Triângulo Mineiro area of Brazil. Rev Soc Bras Med Trop 2018; 51(2): 240-3.
323. Cociancic P, Zonta ML, Navone GT. A cross-sectional study of intestinal parasitoses in dogs and children of the periurban area of La Plata (Buenos Aires, Argentina): Zoonotic importance and implications in public health. Zoonoses Public Health 2018; 65(1): e44-e53.
324. Scanlan PD, Knight R, Song SJ, Ackermann G, Cotter PD. Prevalence and genetic diversity of *Blastocystis* in family units living in the United States. Infect Genet Evol 2016; 45: 95-7.
325. Osman M, Bories J, El Safadi D, *et al.* Prevalence and genetic diversity of the intestinal parasites *Blastocystis* sp. and *Cryptosporidium* spp. in household dogs in France and evaluation of zoonotic transmission risk. Vet Parasitol 2015; 214(1-2): 167-70.
326. Ma L, Qiao H, Wang H, *et al.* Molecular prevalence and subtypes of *Blastocystis* sp. in primates in northern China. Transbound Emerg Dis 2020; 67(6): 2789-96.
327. Crotti D, Sensi M, Crotti S, Grelloni V, Manuali E. *Dientamoeba fragilis* in swine population: a preliminary investigation. Vet Parasitol 2007; 145(3-4): 349-51.

328. Cacciò SM, Sannella AR, Manuali E, *et al.* Pigs as natural hosts of *Dientamoeba fragilis* genotypes found in humans. Emerg Infect Dis 2012; 18(5): 838-41.
329. Mulinge E, Mbae C, Ngugi B, Irungu T, Matey E, Kariuki S. species infection in patients seeking treatment for diarrhea and abdominal discomfort in Mukuru informal settlement in Nairobi, Kenya. Food Waterborne Parasitol 2021; 23: e00122.
330. Santín M, Gómez-Muñoz MT, Solano-Aguilar G, Fayer R. Development of a new PCR protocol to detect and subtype *Blastocystis* spp. from humans and animals. Parasitol Res 2011; 109(1): 205-12.
331. Kilvington S, Price J. Survival of *Legionella pneumophila* within cysts of *Acanthamoeba polyphaga* following chlorine exposure. J Appl Bacteriol 1990; 68(5): 519-25.
332. Kinsella CM, Deijs M, Becker C, *et al.* Host prediction for disease-associated gastrointestinal cressdnaviruses. Virus Evol 2022; 8(2): veac087.
333. Maas L, Dorigo-Zetsma JW, de Groot CJ, Bouter S, Plötz FB, van Ewijk BE. Detection of intestinal protozoa in paediatric patients with gastrointestinal symptoms by multiplex real-time PCR. Clin Microbiol Infect 2014; 20(6): 545-50.
334. Bruijnesteijn van Coppenraet LE, Wallinga JA, Ruijs GJ, Bruins MJ, Verweij JJ. Parasitological diagnosis combining an internally controlled real-time PCR assay for the detection of four protozoa in stool samples with a testing algorithm for microscopy. Clin Microbiol Infect 2009; 15(9): 869-74.
335. Calderaro A, Gorrini C, Montecchini S, *et al.* Evaluation of a real-time polymerase chain reaction assay for the detection of *Dientamoeba fragilis*. Diagn Microbiol Infect Dis 2010; 67(3): 239-45.
336. Stark D, Barratt J, Roberts T, Marriott D, Harkness J, Ellis J. A review of the clinical presentation of dientamoebiasis. Am J Trop Med Hyg 2010; 82(4): 614-9.
337. Stark D, Al-Qassab SE, Barratt JL, *et al.* Evaluation of multiplex tandem real-time PCR for detection of *Cryptosporidium* spp., *Dientamoeba fragilis*, *Entamoeba histolytica*, and *Giardia intestinalis* in clinical stool samples. J Clin Microbiol 2011; 49(1): 257-62.
338. Yildiz I, Erdem Aynur Z. First detection and molecular characterization of *Dientamoeba fragilis* in cattle. Zoonoses Public Health 2022.
339. Yetismis G, Yildirim A, Pekmezci D, *et al.* First report and genotyping of *Dientamoeba fragilis* in pet budgerigars (*Melopsittacus undulatus*), with zoonotic importance. Zoonoses Public Health 2022; 69(5): 572-8.
340. Stark D, Phillips O, Peckett D, *et al.* Gorillas are a host for *Dientamoeba fragilis:* an update on the life cycle and host distribution. Vet Parasitol 2008; 151(1): 21-6.
341. Jinatham V, Popluechai S, Clark CG, Gentekaki E. *Entamoeba chiangraiensis* n. sp. (Amoebozoa: Entamoebidae) isolated from the gut of Asian swamp eel. Parasitology 2019; 146(14): 1719-24.
342. Ai S, Zhang Z, Wang X, Zhang Q, Yin W, Duan Z. The first survey and molecular identification of *Entamoeba* spp. in farm animals on Qinghai-Tibetan Plateau of China. Comp Immunol Microbiol Infect Dis 2021; 75: 101607.
343. Muegge BD, Kuczynski J, Knights D, *et al.* Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. Science 2011; 332(6032): 970-4.
344. Ren M, Yang F, Gou JM, *et al.* First detection and molecular identification of *Entamoeba* in yaks from China. Acta Parasitol 2021; 66(1): 264-70.
345. Levecke B, Dreesen L, Dorny P, *et al.* Molecular identification of *Entamoeba* spp. in captive nonhuman primates. J Clin Microbiol 2010; 48(8): 2988-90.
346. Afonso E, Fu R, Dupaix A, *et al.* Feeding sites promoting wildlife-related tourism might highly expose the endangered Yunnan snub-nosed monkey (*Rhinopithecus bieti*) to parasite transmission. Sci Rep 2021; 11(1): 15817.
347. Villanueva-Garcia C, Gordillo-Chavez EJ, Banos-Ojeda C, *et al.* New *Entamoeba* group in howler monkeys (*Alouatta* spp.) associated with parasites of reptiles. Parasitol Res 2017; 116(8): 2341-6.

348. Garcia G, Ramos F, Perez RG, *et al.* Molecular epidemiology and genetic diversity of *Entamoeba* species in a chelonian collection. J Med Microbiol 2014; 63(Pt 2): 271-83.
349. Tuda J, Feng M, Imada M, Kobayashi S, Cheng X, Tachibana H. Identification of *Entamoeba polecki* with unique 18S rRNA gene sequences from Celebes crested macaques and pigs in Tangkoko Nature Reserve, North Sulawesi, Indonesia. J Eukaryot Microbiol 2016; 63(5): 572-7.
350. Matsubayashi M, Murakoshi N, Komatsu T, Tokoro M, Haritani M, Shibahara T. Genetic identification of *Entamoeba polecki* subtype 3 from pigs in Japan and characterisation of its pathogenic role in ulcerative colitis. Infect Genet Evol 2015; 36: 8-14.
351. Li WC, Geng JZ, Chen C, *et al.* First report on the occurance of intestinal *Entamoeba* spp. in pigs in China. Acta Trop 2018; 185: 385-90.
352. Feng M, Yang B, Yang L, *et al.* High prevalence of *Entamoeba* infections in captive long-tailed macaques in China. Parasitol Res 2011; 109(4): 1093-7.
353. Regan CS, Yon L, Hossain M, Elsheikha HM. Prevalence of *Entamoeba* species in captive primates in zoological gardens in the UK. PeerJ 2014; 2: e492.
354. Ponce Gordo F, Martínez Díaz RA, Herrera S. *Entamoeba struthionis* n.sp. (Sarcomastigophora: Endamoebidae) from ostriches (*Struthio camelus*). Vet Parasitol 2004; 119(4): 327-35.
355. Sylvain PN, Kaur U, Goyal K, Sehgal R, Paul MF. Molecular differentiation of *Entamoeba* spp. isolated from Cameroonian human immunodeficiency virus (HIV) infected and uninfected patient. J Parasitol Vector Biol 2015; 7(7): 139-50.

7. TABLES

Table 1. Single-celled eukaryotic genera observed in humans typically referred to as 'parasites' and that can be found in the human gastrointestinal (GI) tract. Genera highlighted in bold font represent genera for which one or more species are commonly acknowledged as pathogenic.

PROTOZOA				FUNGI	OOMOCYTES
Archamoebae	Sporozoa (Apicomplexa)	Metamonads	Ciliates	Microsporidia	Stramenopiles
Entamoeba*	***Cryptosporidium***	***Giardia***	***Balantioides****	***Enterocytozoon***	*Blastocystis*
Iodamoeba	***Cystoisospora***	*Dientamoeba*		***Encephalitozoon***	
Endolimax	***Cyclospora***	*Chilomastix*			
	Sarcocystis	*Retortamonas*			
		Enteromonas			
		Pentatrichomonas			
		Trichomonas			

*For *Entamoeba*, most species are considered non-pathogenic

**Previously '*Balantidium*' or '*Neobalantidium*'

Table 2. Current guidelines regarding the use of taxonomic terms for studies of genetic diversity of *Entamoeba*. Reproduced from Jacob et al., 2016 [28].

Term	Definition
Species	*Entamoeba* species with Latin binomials have been described primarily on the basis of morphology and host. More recently, some have incorporated molecular data into the species definition. The term "species" is used only where a Latin binomial has been published elsewhere. A sequence identified as belonging to a named species will exhibit a high percentage identity to sequences derived from morphologically verified organisms and will cluster specifically with such sequences to the exclusion of all others with high statistical support.
Subtype (ST)	DNA sequences that cluster as a discrete clade within the range of diversity of a defined species. The identification of a new subtype (ST) must be based on SSU rDNA sequences where gene coverage is ≥80%. Where STs are defined, all sequences within the species must be demarcated into STs. Sequence divergence within a defined ST will not normally be greater than 3%.
Ribosomal lineage (RL)	Organisms for which ≥80% of the SSU rDNA gene has been sequenced, that differ from previously known sequences by 5% or more and where there is no morphological information are assigned ribosomal lineage (RL) numbers. It is possible that RLs could become species in the future if morphological and other relevant data become available, but it might not be appropriate to assign names based only on a single SSU rDNA sequence.
Conditional lineage (CL)	When a divergent sequence does not meet the criteria for a new ribosomal lineage because ≤80% coverage of the SSU rDNA has been obtained, it is proposed identifying it as a conditional lineage (CL). Such lineages are likely to represent novel organisms and to be "promoted" to RLs or species when more data become available.

Table 3. List of consecutive trials of antimicrobial treatment in a Danish patient with irritable bowel syndrome (IBS) with documented *Blastocystis* ST9 and *Dientamoeba fragilis* colonisation (reproduced after [84]).

Agent(s)	Dosage	Clinical effect	Microbiological effect
MZ	750 mg thrice daily/10 days	None	None
Tetracycline	500 mg 4 times daily/10 days	None	None
TMP/SXT	TMP 800 mg/SXT 160 mg twice daily/7 days	None	None
MB + MZ	MB 100 mg thrice separated by 2 weeks, followed by MZ 750 mg thrice daily/10 days	None	None
PM + MZ	PM 500 mg + MZ 750 mg thrice daily/10 days	None	*D. fragilis* eradicated

Abbreviations used: MB, mebendazole; MZ, metronidazole; PM, paromomycin; TMP, trimethoprim; SXT, sulfametaxazole.

Table 4. Examples of differences in positivity rates reported across the globe in studies using DNA-based methods to detect *Dientamoeba fragilis.*

Information	N	Positivity rate (%)	Method	Country	Reference
Stool specimens submitted to SSI	22,000	43	Real-time PCR	Denmark	[144]
Stool specimens from paediatric patients presenting with GI symptoms	163	62	Multiplex qPCR	Netherlands	[333]
Faecal specimens from patients with GI complaints	397	32	Combination of LM and qPCR	Netherlands	[334]
Faecal specimens from patients with clinical suspicion of intestinal parasitosis	497	21	qPCR	Italy	[335]
Faecal specimens from patients with GI complaints	750	5.2	qPCR	Australia	[336]
Faecal specimens submitted to the Department of Microbiology at St. Vincent's Hospital, Sydney	472	5.5	Tandem Multiplex PCR	Australia	[337]

Table 5. Overview of *Dientamoeba fragilis*-specific SSU rDNA sequences (18S only; not ITS sequences with flanking regions generated by Windsor and colleagues) currently available in the NCBI nucleotide database (January 2023) listed according to date of deposition.

NCBI Acc. no.	Organism	Genotype	Host	Isolation source	Country	Sequence length	Date of deposition
MZ405082	*Dientamoeba fragilis*	1	*Homo sapiens*	stool	Turkey	361	24-APR-2022
ON242172	*Dientamoeba fragilis*	1	NA	NA	NA	810	20-APR-2022
OM250406	*Dientamoeba fragilis*	1	*Homo sapiens*	faeces	Turkey	788	18-JAN-2022
MW130447	*Dientamoeba fragilis*	1	Budgerigar	faeces	Turkey	324	24-OCT-2020
MW130448	*Dientamoeba fragilis*	1	Budgerigar	faeces	Turkey	324	24-OCT-2020
MN914083	*Dientamoeba fragilis*	1	*Homo sapiens*	faeces	Germany	1621	07-JUL-2020
MN920432	*Dientamoeba fragilis*	1	*Homo sapiens*	NA	Brazil	184	15-JAN-2020
MN920434	*Dientamoeba fragilis*	1	*Homo sapiens*	NA	Brazil	184	15-JAN-2020
MN920433	*Dientamoeba fragilis*	1	*Homo sapiens*	NA	Brazil	184	15-JAN-2020
MN920435	*Dientamoeba fragilis*	1	*Homo sapiens*	NA	Brazil	184	15-JAN-2020
MN920436	*Dientamoeba fragilis*	1	*Homo sapiens*	NA	Brazil	184	15-JAN-2020
MN920437	*Dientamoeba fragilis*	1	*Homo sapiens*	NA	Brazil	184	15-JAN-2020
MN920438	*Dientamoeba fragilis*	1	*Homo sapiens*	NA	Brazil	184	15-JAN-2020
MN920439	*Dientamoeba fragilis*	1	*Homo sapiens*	NA	Brazil	184	15-JAN-2020
MN560149	*Dientamoeba fragilis*	1	*Homo sapiens*	faeces	Turkey	806	20-OCT-2019
MN560150	*Dientamoeba fragilis*	1	*Homo sapiens*	faeces	Turkey	806	20-OCT-2019
MN183755	*Dientamoeba fragilis*	1	NA	NA	NA	188	22-JUL-2019
MN183756	*Dientamoeba fragilis*	1	NA	NA	NA	188	22-JUL-2019
MN183757	*Dientamoeba fragilis*	1	NA	NA	NA	188	22-JUL-2019
MN183758	*Dientamoeba fragilis*	1	NA	NA	NA	188	22-JUL-2019
MN183759	*Dientamoeba fragilis*	1	NA	NA	NA	188	22-JUL-2019
MN183761	*Dientamoeba fragilis*	1	NA	NA	NA	188	22-JUL-2019
MN183760	*Dientamoeba fragilis*	1	NA	NA	NA	188	22-JUL-2019
MN183762	*Dientamoeba fragilis*	1	NA	NA	NA	188	22-JUL-2019
MN183763	*Dientamoeba fragilis*	1	NA	NA	NA	188	22-JUL-2019
MN183764	*Dientamoeba fragilis*	1	NA	NA	NA	188	22-JUL-2019
MN183765	*Dientamoeba fragilis*	1	NA	NA	NA	188	22-JUL-2019

MN183766	*Dientamoeba fragilis*	1	NA	NA	NA	187	22-JUL-2019
MN183767	*Dientamoeba fragilis*	1	NA	NA	NA	187	22-JUL-2019
AB692771	*Dientamoeba fragilis*	1	*Homo sapiens*	NA	Iran	842	17-JAN-2012
AB692772	*Dientamoeba fragilis*	1	*Homo sapiens*	NA	Iran	856	17-JAN-2012
AB692773	*Dientamoeba fragilis*	1	*Homo sapiens*	NA	Iran	660	17-JAN-2012
KU939320	*Dientamoeba fragilis*	1	NA	NA	NA	317	12-JUN-2016
JQ677147	*Dientamoeba fragilis*	1	*Homo sapiens*	faeces	United Kingdom	1501	19-MAY-2012
JQ677148	*Dientamoeba fragilis*	1	*Homo sapiens*	faeces	United Kingdom	1501	19-MAY-2012
JQ677149	*Dientamoeba fragilis*	1	*Homo sapiens*	faeces	Italy	1085	19-MAY-2012
JQ677150	*Dientamoeba fragilis*	1	*Homo sapiens*	faeces	Italy	355	19-MAY-2012
JQ677151	*Dientamoeba fragilis*	1	*Homo sapiens*	faeces	Italy	355	19-MAY-2012
JQ677152	*Dientamoeba fragilis*	1	*Homo sapiens*	faeces	Italy	355	19-MAY-2012
JQ677153	*Dientamoeba fragilis*	1	*Homo sapiens*	faeces	Italy	355	19-MAY-2012
JQ677154	*Dientamoeba fragilis*	1	*Homo sapiens*	faeces	Italy	355	19-MAY-2012
JQ677155	*Dientamoeba fragilis*	1	*Homo sapiens*	faeces	Italy	355	19-MAY-2012
JQ677156	*Dientamoeba fragilis*	1	*Homo sapiens*	faeces	Italy	355	19-MAY-2012
JQ677157	*Dientamoeba fragilis*	1	pig	faeces	Italy	355	19-MAY-2012
JQ677158	*Dientamoeba fragilis*	1	pig	faeces	Italy	355	19-MAY-2012
JQ677159	*Dientamoeba fragilis*	1	pig	faeces	Italy	355	19-MAY-2012
JQ677160	*Dientamoeba fragilis*	1	pig	faeces	Italy	355	19-MAY-2012
JQ677161	*Dientamoeba fragilis*	1	pig	faeces	Italy	355	19-MAY-2012
JQ677162	*Dientamoeba fragilis*	1	pig	faeces	Italy	355	19-MAY-2012
FJ649228	*Dientamoeba fragilis*	1	NA	NA	NA	792	23-FEB-2009
AY730405	*Dientamoeba fragilis*	1	NA	NA	Australia	1661	28-SEP-2004
U37461	*Dientamoeba fragilis*	2	NA	NA	NA	1676	11-JUL-1996

Table 6. Data from studies involving DNA-based screening of faecal samples from non-human hosts for *Dientamoeba fragilis.*

Host species	Number of individuals tested	Country of study	Positivity rate	Genotype (if available)	Reference
Pig (*Sus scrofa*)	152	Italy	47%	1	[328]
Pig	37	Denmark	0%	NA	[13]
Cattle	163	Turkey	19%	1	[338]
Pet budgerigar	150	Turkey	21.3%	1	[339]
Western low-land gorillas	10	Australia (zoo)	30%	NA	[340]*

*In the study by Stark and colleagues, samples were obtained from wild birds (n = 79), primates (n = 45) from several species collected from Taronga Zoo, Sydney, swine (n = 135), cows (n = 50), sheep (50), horses (25) and goats (n = 25) and domestic animals including dogs (n = 50), cats (n = 50), mice (n = 25), rats (n = 25), guinea pigs (n = 20), rabbits (n = 20), and other mammals (n = 9). *Dientamoeba fragilis* was found only in gorillas.
NA, not applicable.

Table 7. Overview of species and ribosomal lineages (RL) of *Entamoeba* based on DNA sequence data available in the NCBI nucleotide database (update of the table included in book chapter by Clark and Stensvold [181] based mainly on references [28, 44, 181, 341]). For more detailed information on ribosomal lineages and conditional lineages, see tables 2 and 8. In this table, 'Ungulates' include elephants for practical reasons.

				HOST									
Species/lineage	Subtype	Potentially Invasive[#]	Number of cyst nuclei	Environment	Humans	NHP	Ungulates	Rodents	Carnivores	Birds	Reptiles	Amphibia	Fish
E. histolytica	-	X	4	-	X	X	X	-	X	-	-	-	-
E. dispar	-	-	4	-	X	X	-	-	X?	-	-	-	-
E. bangladeshi	-	-	4	-	X	-	-	-	-	-	-	-	-
E. moshkovskii	-	-	4	X	X	-	X[¥]	-	X?	-	X	-	-
E. nuttalli	-	X	4	-	X[¥¥]	X	-	-	-	-	-	-	-
E. ecuadoriensis	-	-	4	X	-	-	-	-	-	-	-	-	-
E. bovis	-	-	1	-	-	-	X	-	-	-	-	-	-
Entamoeba RL1	-	-	1	-	-	-	X	-	-	-	-	-	-
Entamoeba RL2	-	-	NA*	-	-	-	X	-	-	-	-	-	-
Entamoeba RL3	-	-	1	-	-	X	X	-	-	-	-	-	-
Entamobea RL4	-	-	NA	-	-	-	X	-	-	-	-	-	-
Entamoeba RL5	-	-	4	-	-	-	-	-	-	-	X	-	-
Entamoeba RL6	-	-	4	-	-	-	-	-	-	-	X	-	-
Entamoeba RL8	-	-	NA	-	-	-	X	-	-	-	-	-	-
Entamoeba RL9	-	-	NA	-	-	-	X	-	-	-	-	-	-
E. terrapinae	-	-	4	-	-	-	-	-	-	-	X	-	-
E. insolita	-	-	4	-	-	-	-	-	-	-	X	-	-
E. hartmanni	ST1	-	4	-	X	X	-	-	X	-	-	-	-
	ST2	-	4	-	X	-	-	-	-	-	-	-	-
	ST3	-	4	-	X	-	-	-	-	-	-	-	-
Entamoeba RL10	-	-	NA	-	-	-	X	-	-	-	-	-	-

E. equi	-	-	NA	-	-	-	X	-	-	-	-	-	-
E. ranarum	-	-	1	-	-	-	-	-	-	-	X	X	-
E. invadens	-	X	1	-	-	-	-	-	-	-	X	-	-
E. chiangraiensis	-	-	1	-	-	-	-	-	-	-	-	-	X
E. suis	-	-	1	-	-	X	X	-	-	-	-	-	-
E. marina	-	-	4	X	-	-	-	-	-	-	-	-	-
*E. gingivalis***	ST1	-	None	-	X	-	-	-	-	-	-	-	-
	ST2	-	None	-	X	-	-	-	-	-	-	-	-
	ST3	-	None	-	X								
E. polecki	ST1	-	1	-	X	-	X	-	X	X	-	-	-
	ST2	-	1	-	X	X	-	-	-	-	-	-	-
	ST3	-	1	-	X	-	X	-	-	X	-	-	-
	ST4	-	1	-	X	X	-	-	-	-	-	-	-
Entamoeba RL7	-	-	8**	-	X	X	X***	-	-	-	-	-	-
Entamoeba RL11	-	-	NA	-	-	-	-	X	-	-	-	-	-
E. muris	-	-	8	-	-	-	-	X	-	-	-	-	-
E. coli	ST1	-	8	-	X	X	-	-	-	-	-	-	-
	ST2	-	8	-	X	X	-	X	-	-	-	-	-
	ST3	-	8	-	X	-	-	-	-	-	-	-	-

[#]only those species for which invasiveness has been demonstrated are marked as being potentially invasive
*possibly 4 nuclei per cyst [49]
**probably 8 nuclei per cyst (Vidal-Lapiedra, unpublished observations)
***Lebbad et al., unpublished observations (1,306 bp sequence with 99% identity to RL7)
****_E. gingivalis_ has been found at least once in a carnivorous host (dog) [237]
[¥]Found in pigs by nested PCR + sequencing [247] and in horses by single-round PCR + sequencing [342]
[¥¥]Found once in an animal caretaker working in a zoo [178]
NA: Information not available (no morphology data available)

Table 8. *Entamoeba* ribosomal lineages (RL) and conditional lineages (CL) reported until January 2023.

Lineage name	Host species	Geographical information	NCBI Acc. no.	Reference
RL1	*Gazella spekei*	US - Saint Louis Zoo	NA	[28, 306, 343]
	Ovis canadensis	US - Saint Louis Zoo	NA	[28, 306, 343]
	Capreolus capreolus	Sweden	FN666253	[50]
RL2	*Bos taurus*	Sweden	FR686362	[49]
	Bos taurus	Sweden	FR686363	[49]
	Ovis aries	UK		[28]
	Bos grunniens	Tibet	MH127446	[344]
	Capra hircus	Iraq	MF568375	Allban, unpublished
	Bos taurus	China	KC922175	Zhang, unpublished
RL3	*Trachypithecus auratus* or *T. cristatus*	UK – zoo	FR686359	[49]
	Semnopithecus entellus	Germany – Zoologischer Garten Neunkirchen	FR686358	[49]
	Semnopithecus entellus	Belgium - Zoo	GU437826	[345]
RL4	*Bos taurus*	Libya	FR686361	[49]
	Bos taurus	Estonia	FR686451	[49]
	Bos taurus	UK		[28]
	Bos taurus	Tibet	NA	[346]
RL5	*Psammobates pardalis*	Sweden	FR686365	[49]
RL6	*Iguana iguana*	US - National Zoo, Washington DC	AF149911	[30, 49]
	Allouata pigra	Mexico	KY620097	[347]
RL7	*Trachypithecus phayrei*	UK – zoo	FR686360	[49]
RL8	*Bos taurus*	UK	KR025406	[28]
	Gazella spekei	US - Saint Louis Zoo	NA	[28, 306, 343]
	Okapia johnstoni	US - Saint Louis Zoo	NA	[28, 306, 343]
	Bos taurus	UK		[28]
	Camel	China	MN749974	[342]
	Capra hircus	Iraq	MF568373	Allban, unpublished
RL9	*Equus ferus caballus*	UK	KR025407	[28]
	Bos grunniens	Tibet	MH127441	[344]
	'Horse'	China	MN749978	[342]

RL10	*Elephas maximus*	The Netherlands - Zoo	KR025408	[28]
RL11	*Myodes glareolus*	UK	KR025409	[28]
CL1	*Aldabrachelys gigantea*	Mauritius	KR025410	[28]
CL2	'Aquatic turtles'	Mexico	JQ406871	[28, 348]
CL3	*Okapia johnstoni*	US - Saint Louis Zoo	NA	[28, 306, 343]
CL4	*Okapia johnstoni*	US - Saint Louis Zoo	NA	[28, 306, 343]
CL5	*Okapia johnstoni*	US - Saint Louis Zoo	NA	[28, 306, 343]
	Ovis canadensis	US - Saint Louis Zoo	NA	[28, 306, 343]
CL6	*Equus acinus*	US - Saint Louis Zoo	NA	[28, 306, 343]
CL7	Kangaroo?	Unknown	NA	[28, 306]
CL8	*Alluata pigra*	Mexico	KY620092	[347]
	Alluata palliata	Mexico	KY620090	[347]
CL9	*Equus* sp.	Tibet	MW718196	[346]
CL10	*Bos taurus*	Tibet	MW718202	[346]
CL11	'Dog'	India	OP919601	Jayasri, unpublished

Table 9. *Entamoeba coli*-specific SSU rDNA sequences currently available in the NCBI nucleotide database (January 2023) listed according to date of deposition. Subtype information has been provided where possible. A few sequences in the database that cluster with *E. coli* are listed as '*Entamoeba* sp.' and not included in this overview; these include, but may not be limited to, OP796187-OP796190, FN396614, MG256518, MG256520, MG256528, MW133765, MW133768, KY658155, KY658156, KY658157, KY658172, KY658178 and KY658179). Moreover FN396613 and FN396614 are listed in GenBank as *Entamoeba muris*; however, these are likely *Entamoeba coli* ST2.

NCBI Acc. no.	Organism	Subtype	Host	Isolation source	Country	Sequence length	Date of deposition
OQ241737	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	Argentina	583	12-JAN-2023
OP868730	*Entamoeba coli*	ST3	*Homo sapiens*	NA	Iraq	420	26-NOV-2022
OP868731	*Entamoeba coli*	ST3	*Homo sapiens*	NA	Iraq	420	26-NOV-2022
OP868732	*Entamoeba coli*	ST3	*Homo sapiens*	NA	Iraq	420	26-NOV-2022
ON713469	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	Argentina	580	31-OCT-2022
ON989959	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	China	664	20-JUL-2022
ON989960	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	China	664	20-JUL-2022
ON989961	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989962	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989963	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989964	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989965	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989966	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989967	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989968	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989969	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989970	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989971	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989972	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989973	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989974	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	659	20-JUL-2022
ON989975	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989976	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	659	20-JUL-2022
ON989977	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989978	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022

ON989979	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	664	20-JUL-2022
ON989980	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989981	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	664	20-JUL-2022
ON989982	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	661	20-JUL-2022
ON989983	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	661	20-JUL-2022
ON989984	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	664	20-JUL-2022
ON989985	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	664	20-JUL-2022
ON989986	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989987	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989988	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989989	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	664	20-JUL-2022
ON989990	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989991	*Entamoeba coli*	ST3	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989992	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	664	20-JUL-2022
ON989993	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	664	20-JUL-2022
ON989994	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	664	20-JUL-2022
ON989995	*Entamoeba coli*	ST3	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989996	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	659	20-JUL-2022
ON989997	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON989998	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	664	20-JUL-2022
ON989999	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990000	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	664	20-JUL-2022
ON990001	*Entamoeba coli*	ST3	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990002	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	664	20-JUL-2022
ON990003	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990004	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	664	20-JUL-2022
ON990005	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	664	20-JUL-2022
ON990006	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	664	20-JUL-2022
ON990007	*Entamoeba coli*	ST3	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990008	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022

ON990009	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	664	20-JUL-2022
ON990010	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990011	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	664	20-JUL-2022
ON990012	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	664	20-JUL-2022
ON990013	*Entamoeba coli*	ST3	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990014	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990015	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990016	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990017	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990018	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990019	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990020	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990021	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990022	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990023	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990024	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990025	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990026	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990027	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990028	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990029	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	659	20-JUL-2022
ON990030	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	664	20-JUL-2022
ON990031	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990032	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990033	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990034	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	NA	665	20-JUL-2022
ON990035	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	NA	664	20-JUL-2022
ON254801	*Entamoeba coli*	Possibly not *E. coli*	NHP	NA	China	125	29-JUN-2022
ON254802	*Entamoeba coli*	ST2	NHP	NA	China	129	29-JUN-2022
MW819961	*Entamoeba coli*	ST2	*Macaca thibetana*	NA	China	592	28-APR-2022

OM985616	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	Argentina	565	19-MAR-2022
OM985617	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	Argentina	560	19-MAR-2022
OM985619	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	Argentina	603	19-MAR-2022
OM985620	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	Argentina	572	19-MAR-2022
MZ787759	*Entamoeba coli*	ST1	*Homo sapiens*	faecal samples	NA	583	18-AUG-2021
MZ787760	*Entamoeba coli*	ST1	*Homo sapiens*	faecal samples	NA	568	18-AUG-2021
MW026735	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026736	*Entamoeba coli*	ST3	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026737	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	Brazil	528	30-MAR-2021
MW026738	*Entamoeba coli*	unresolved	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026739	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026740	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	Brazil	528	30-MAR-2021
MW026741	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026742	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	Brazil	528	30-MAR-2021
MW026743	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	Brazil	528	30-MAR-2021
MW026744	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	Brazil	528	30-MAR-2021
MW026745	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	Brazil	526	30-MAR-2021
MW026746	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026747	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026748	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026749	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026750	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026751	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026752	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026753	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026754	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026755	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026756	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026757	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026758	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021

MW026759	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026760	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026761	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026762	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026763	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026764	*Entamoeba coli*	ST2	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026765	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MW026766	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	Brazil	527	30-MAR-2021
MN914079	*Entamoeba coli*	ST1	*Homo sapiens*	human faeces	Germany	1174	07-JUL-2020
MK541024	*Entamoeba coli*	ST1	*Homo sapiens*	faeces	Argentina	631	03-APR-2019
MK541025	*Entamoeba coli*	ST1	NA	faeces	Mexico	601	03-APR-2019
MH133210	*Entamoeba coli*	ST2	rhesus macaques	faecal samples	China	630	01-APR-2019
MH623050	*Entamoeba coli*	ST1	environmental	waste water	Australia	107	24-MAR-2019
MK559460	*Entamoeba coli*	ST3	rhesus monkeys	faeces	NA	154	27-FEB-2019
MK559461	*Entamoeba coli*	ST3	rhesus monkeys	faeces	NA	160	27-FEB-2019
MK559462	*Entamoeba coli*	ST2	rhesus monkeys	faeces	NA	160	27-FEB-2019
MH629959	*Entamoeba coli*	ST1	*Homo sapiens*	faecal matter	West Bank	437	22-JUL-2018
MH629962	*Entamoeba coli*	ST2	*Homo sapiens*	faecal matter	West Bank	429	22-JUL-2018
MH629963	*Entamoeba coli*	ST2	*Homo sapiens*	faecal matter	West Bank	429	22-JUL-2018
MH629964	*Entamoeba coli*	ST3	*Homo sapiens*	faecal matter	West Bank	433	22-JUL-2018
MH629965	*Entamoeba coli*	ST1	*Homo sapiens*	faecal matter	West Bank	438	22-JUL-2018
MH629966	*Entamoeba coli*	ST1	*Homo sapiens*	faecal matter	West Bank	437	22-JUL-2018
MH620469	*Entamoeba coli*	ST2	NA	NA	NA	566	20-JUL-2018
MH620403	*Entamoeba coli*	ST2	NA	NA	NA	604	19-JUL-2018
MF631996	*Entamoeba coli*	ST1	*Homo sapiens*	xenic culture	NA	182	28-MAR-2018
MF631997	*Entamoeba coli*	ST2	*Gorilla gorilla*	xenic culture	NA	220	28-MAR-2018
MG925061	*Entamoeba coli*	ST1	*Homo sapiens*	stool	Tunisia	539	14-FEB-2018
MG925062	*Entamoeba coli*	ST1	*Homo sapiens*	stool	Tunisia	539	14-FEB-2018
MG925063	*Entamoeba coli*	ST3	*Homo sapiens*	stool	Tunisia	536	14-FEB-2018
MG925064	*Entamoeba coli*	ST2	*Homo sapiens*	stool	Tunisia	531	14-FEB-2018

KX923799	*Entamoeba coli*	ST2	NHP	NA	China	127	23-SEP-2017
KX923800	*Entamoeba coli*	ST2	NHP	NA	China	126	23-SEP-2017
KX923801	*Entamoeba coli*	ST2	NHP	NA	China	126	23-SEP-2017
KX923802	*Entamoeba coli*	ST2	NHP	NA	China	126	23-SEP-2017
KU320610	*Entamoeba coli*	ST2	*Pan troglodytes schweinfurthii*	NA	NA	587	13-MAR-2016
KU320611	*Entamoeba coli*	ST2	*Pan troglodytes schweinfurthii*	NA	NA	587	13-MAR-2016
AB851495	*Entamoeba coli*	ST2	*Homo sapiens*	HIV patient	Cameroon	358	23-FEB-2016
AB851497	*Entamoeba coli*	ST2	*Homo sapiens*	HIV patient	Cameroon	354	23-FEB-2016
AB851499	*Entamoeba coli*	ST2	*Homo sapiens*	HIV patient	Cameroon	329	23-FEB-2016
AB851500	*Entamoeba coli*	ST2	*Homo sapiens*	HIV patient	Cameroon	331	23-FEB-2016
AB845674	*Entamoeba coli*	Not *Entamoeba*	*Homo sapiens*	HIV patient	Cameroon	313	16-JAN-2016
AB444953	*Entamoeba coli*	ST2	*Gorilla gorilla*	faeces	Japan	2106	02-JUN-2009
FR686364	*Entamoeba coli*	ST3	*Homo sapiens*	host faeces	Nigeria	2100	18-OCT-2010
FR686401	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	506	18-OCT-2010
FR686402	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1260	18-OCT-2010
FR686403	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1260	18-OCT-2010
FR686404	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1261	18-OCT-2010
FR686405	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	271	18-OCT-2010
FR686406	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1260	18-OCT-2010
FR686407	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1261	18-OCT-2010
FR686408	*Entamoeba coli*	ST2	*Homo sapiens*	host faeces	NA	1262	18-OCT-2010
FR686409	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	460	18-OCT-2010
FR686410	*Entamoeba coli*	ST1	*Mandrillus leucophaeus*	host faeces	Germany	582	18-OCT-2010
FR686411	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1379	18-OCT-2010
FR686412	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1375	18-OCT-2010
FR686413	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1377	18-OCT-2010
FR686414	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1260	18-OCT-2010
FR686415	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1380	18-OCT-2010
FR686416	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1260	18-OCT-2010
FR686417	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1261	18-OCT-2010

FR686418	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1377	18-OCT-2010
FR686419	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1261	18-OCT-2010
FR686420	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1261	18-OCT-2010
FR686421	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1380	18-OCT-2010
FR686422	*Entamoeba coli*	ST2	*Homo sapiens*	host faeces	NA	1262	18-OCT-2010
FR686423	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	Brazil	1260	18-OCT-2010
FR686424	*Entamoeba coli*	ST3	*Homo sapiens*	host faeces	NA	163	18-OCT-2010
FR686425	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	Cyprus	1261	18-OCT-2010
FR686426	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1260	18-OCT-2010
FR686427	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1261	18-OCT-2010
FR686428	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	Lebanon	1261	18-OCT-2010
FR686429	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1261	18-OCT-2010
FR686430	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1261	18-OCT-2010
FR686431	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1261	18-OCT-2010
FR686432	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1261	18-OCT-2010
FR686433	*Entamoeba coli*	ST2	*Homo sapiens*	host faeces	Viet Nam	583	18-OCT-2010
FR686434	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	931	18-OCT-2010
FR686435	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1260	18-OCT-2010
FR686436	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1261	18-OCT-2010
FR686437	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	1261	18-OCT-2010
FR686438	*Entamoeba coli*	ST1	*Homo sapiens*	host faeces	NA	845	18-OCT-2010
FR686439	*Entamoeba coli*	ST2	*Chinchilla* sp.	host faeces	Belgium	938	18-OCT-2010
FR686440	*Entamoeba coli*	ST2	*Homo sapiens*	host faeces	NA	1371	18-OCT-2010
FR686441	*Entamoeba coli*	ST2	*Homo sapiens*	host faeces	NA	1262	18-OCT-2010
FR686442	*Entamoeba coli*	ST2	*Homo sapiens*	host faeces	Rwanda	1055	18-OCT-2010
FR686443	*Entamoeba coli*	ST2	*Homo sapiens*	host faeces	Peru	1262	18-OCT-2010
FR686444	*Entamoeba coli*	ST2	*Homo sapiens*	host faeces	Tanzania	1322	18-OCT-2010
FR686445	*Entamoeba coli*	ST2	*Homo sapiens*	host faeces	Malawi	1262	18-OCT-2010
FR686446	*Entamoeba coli*	ST2	*Homo sapiens*	host faeces	Ecuador	1261	18-OCT-2010
FR686447	*Entamoeba coli*	ST2	*Gorilla gorilla*	host faeces	Germany	583	18-OCT-2010

FR686448	*Entamoeba coli*	ST2	*Homo sapiens*	host faeces	Viet Nam	583	18-OCT-2010
FR686449	*Entamoeba coli*	ST2	*Macaca fuscata*	host faeces	NA	806	18-OCT-2010
KF646245	*Entamoeba coli*	ST2	NA	NA	NA	406	30-MAR-2015
KF646246	*Entamoeba coli*	ST2	NA	NA	NA	362	30-MAR-2015
KF646247	*Entamoeba coli*	ST2	NA	NA	NA	509	30-MAR-2015
AB749456	*Entamoeba coli*	ST2	*Macaca mulata*	NA	China	111	22-SEP-2012
AB749457	*Entamoeba coli*	ST1	*Macaca fascicularis*	NA	China	110	22-SEP-2012
AB749458	*Entamoeba coli*	ST2	*Macaca fascicularis*	NA	China	112	22-SEP-2012
AF149914	*Entamoeba coli*	ST2	NA	NA	NA	2101	22-DEC-1999
AF149915	*Entamoeba coli*	ST1	NA	NA	NA	2104	22-DEC-1999

Table 10. *Entamoeba dispar*-specific SSU rDNA sequences currently available in the NCBI nucleotide database (January 2023) listed according to date of deposition.

NCBI Acc. no.	Organism	Host	Isolation source	Country	Sequence length	Date of deposition
OP874659	*Entamoeba dispar*	*Homo sapiens*	stool	NA	573	27-NOV-2022
OP874660	*Entamoeba dispar*	*Homo sapiens*	stool	NA	573	27-NOV-2022
OP874661	*Entamoeba dispar*	*Homo sapiens*	stool	NA	573	27-NOV-2022
OP874662	*Entamoeba dispar*	*Homo sapiens*	stool	NA	573	27-NOV-2022
OP874663	*Entamoeba dispar*	*Homo sapiens*	stool	NA	573	27-NOV-2022
OP874688	*Entamoeba dispar*	*Homo sapiens*	stool	NA	573	27-NOV-2022
OP874689	*Entamoeba dispar*	*Homo sapiens*	stool	NA	573	27-NOV-2022
OP874690	*Entamoeba dispar*	*Homo sapiens*	stool	NA	560	27-NOV-2022
OP874691	*Entamoeba dispar*	*Homo sapiens*	stool	NA	573	27-NOV-2022
OP874692	*Entamoeba dispar*	*Homo sapiens*	stool	NA	586	27-NOV-2022
ON668115	*Entamoeba dispar*	*Homo sapiens*	faeces	Colombia	177	14-OCT-2022
OP524421	*Entamoeba dispar*	*Homo sapiens*	stool	Iraq	1746	01-OCT-2022
OP524422	*Entamoeba dispar*	*Homo sapiens*	stool	Iraq	1728	01-OCT-2022
OP524423	*Entamoeba dispar*	*Homo sapiens*	stool	Iraq	1792	01-OCT-2022
OP524424	*Entamoeba dispar*	*Homo sapiens*	stool	Iraq	1714	01-OCT-2022
OP524425	*Entamoeba dispar*	*Homo sapiens*	stool	Iraq	1727	01-OCT-2022
OP451882	*Entamoeba dispar*	*Homo sapiens*	stool	NA	468	21-SEP-2022
OP451883	*Entamoeba dispar*	*Homo sapiens*	stool	NA	461	21-SEP-2022
OP451884	*Entamoeba dispar*	*Homo sapiens*	stool	NA	178	21-SEP-2022
OP451885	*Entamoeba dispar*	*Homo sapiens*	stool	NA	159	21-SEP-2022
OP451886	*Entamoeba dispar*	*Homo sapiens*	stool	NA	148	21-SEP-2022
OP453103	*Entamoeba dispar*	Human and Domestic dogs	stool	Iraq	468	21-SEP-2022
OP453104	*Entamoeba dispar*	Human and Domestic dogs	stool	Iraq	461	21-SEP-2022
OP453105	*Entamoeba dispar*	Human and Domestic dogs	stool	Iraq	178	21-SEP-2022
OP453106	*Entamoeba dispar*	Human and Domestic dogs	stool	Iraq	1714	21-SEP-2022
OP453107	*Entamoeba dispar*	Human and Domestic dogs	stool	Iraq	318	21-SEP-2022
ON254803	*Entamoeba dispar*	NHPs	NA	China	129	29-JUN-2022
ON692812	*Entamoeba dispar*	*Homo sapiens*	stool	Iraq	166	12-JUN-2022

ON318865	*Entamoeba dispar*	NA	faeces	Malaysia	174	27-APR-2022
ON318866	*Entamoeba dispar*	NA	faeces	Malaysia	174	27-APR-2022
ON318867	*Entamoeba dispar*	NA	faeces	Malaysia	174	27-APR-2022
OM985615	*Entamoeba dispar*	*Homo sapiens*	faeces	Argentina	510	19-MAR-2022
OM985618	*Entamoeba dispar*	*Homo sapiens*	faeces	Argentina	511	19-MAR-2022
OM791700	*Entamoeba dispar*	*Homo sapiens*	stool	India	1147	28-FEB-2022
OM791701	*Entamoeba dispar*	*Homo sapiens*	stool	India	1147	28-FEB-2022
OM268856	*Entamoeba dispar*	*Homo sapiens*	stool	Iraq	461	19-JAN-2022
OM268859	*Entamoeba dispar*	*Homo sapiens*	stool	Iraq	468	19-JAN-2022
OM190405	*Entamoeba dispar*	*Homo sapiens*	stool	Iran	133	15-JAN-2022
MZ913017	*Entamoeba dispar*	Human, Child	stool	Iraq	159	30-AUG-2021
MZ913023	*Entamoeba dispar*	Human, Child	stool	Iraq	158	30-AUG-2021
MZ787761	*Entamoeba dispar*	*Homo sapiens*	faecal samples		491	18-AUG-2021
MW026767	*Entamoeba dispar*	*Homo sapiens*	faeces	Brazil	483	30-MAR-2021
MW026768	*Entamoeba dispar*	*Homo sapiens*	faeces	Brazil	484	30-MAR-2021
MW026769	*Entamoeba dispar*	*Homo sapiens*	faeces	Brazil	484	30-MAR-2021
MW026770	*Entamoeba dispar*	*Homo sapiens*	faeces	Brazil	483	30-MAR-2021
MW026771	*Entamoeba dispar*	*Homo sapiens*	faeces	Brazil	483	30-MAR-2021
MW026772	*Entamoeba dispar*	*Homo sapiens*	faeces	Brazil	483	30-MAR-2021
MW026773	*Entamoeba dispar*	*Homo sapiens*	faeces	Brazil	483	30-MAR-2021
MW026774	*Entamoeba dispar*	*Homo sapiens*	faeces	Brazil	483	30-MAR-2021
MW026775	*Entamoeba dispar*	*Homo sapiens*	faeces	Brazil	483	30-MAR-2021
MW026776	*Entamoeba dispar*	*Homo sapiens*	faeces	Brazil	483	30-MAR-2021
MW026777	*Entamoeba dispar*	*Homo sapiens*	faeces	Brazil	483	30-MAR-2021
MW026778	*Entamoeba dispar*	*Homo sapiens*	faeces	Brazil	483	30-MAR-2021
MW026779	*Entamoeba dispar*	*Homo sapiens*	faeces	Brazil	483	30-MAR-2021
MW026780	*Entamoeba dispar*	*Homo sapiens*	faeces	Brazil	483	30-MAR-2021
MW026781	*Entamoeba dispar*	*Homo sapiens*	faeces	Brazil	482	30-MAR-2021
MW026782	*Entamoeba dispar*	*Homo sapiens*	faeces	Brazil	483	30-MAR-2021
MW026783	*Entamoeba dispar*	*Homo sapiens*	faeces	Brazil	483	30-MAR-2021
MW026784	*Entamoeba dispar*	*Homo sapiens*	faeces	Brazil	483	30-MAR-2021

MW624412	*Entamoeba dispar*	NA	NA	NA	169	23-FEB-2021
MW624413	*Entamoeba dispar*	NA	NA	NA	169	23-FEB-2021
MW624414	*Entamoeba dispar*	NA	NA	NA	169	23-FEB-2021
MW165339	*Entamoeba dispar*	*Homo sapiens*	stool	Iraq	148	28-OCT-2020
MW029816	*Entamoeba dispar*	*Homo sapiens*	stool	Iraq	308	27-SEP-2020
MW029817	*Entamoeba dispar*	*Homo sapiens*	stool	Iraq	318	27-SEP-2020
MT250839	*Entamoeba dispar*	*Homo sapiens*	NA	Iraq	178	31-MAR-2020
MH754938	*Entamoeba dispar*	*Homo sapiens*	NA	Kenya	177	06-SEP-2019
MH754939	*Entamoeba dispar*	*Homo sapiens*	NA	Kenya	276	06-SEP-2019
MK541026	*Entamoeba dispar*	*Homo sapiens*	faeces	Argentina	591	03-APR-2019
MH133211	*Entamoeba dispar*	rhesus macaques	faecal samples	China	395	01-APR-2019
MK559463	*Entamoeba dispar*	rhesus monkeys	faeces	NA	755	27-FEB-2019
MK559464	*Entamoeba dispar*	rhesus monkeys	faeces	NA	756	27-FEB-2019
MK559465	*Entamoeba dispar*	rhesus monkeys	faeces	NA	833	27-FEB-2019
MH629960	*Entamoeba dispar*	*Homo sapiens*	faecal-matter	West Bank	365	22-JUL-2018
MH629961	*Entamoeba dispar*	*Homo sapiens*	faecal-matter	West Bank	366	22-JUL-2018
MF631990	*Entamoeba dispar*	NA	axenic culture	NA	254	28-MAR-2018
MG256516	*Entamoeba dispar*	*Homo sapiens*	faecal sample	NA	227	28-MAR-2018
KY823418	*Entamoeba dispar*	NA	NA	NA	136	26-FEB-2018
KY823419	*Entamoeba dispar*	NA	NA	NA	139	26-FEB-2018
KY823420	*Entamoeba dispar*	NA	NA	NA	137	26-FEB-2018
KY823421	*Entamoeba dispar*	NA	NA	NA	128	26-FEB-2018
KY823422	*Entamoeba dispar*	NA	NA	NA	129	26-FEB-2018
KY823423	*Entamoeba dispar*	NA	NA	NA	133	26-FEB-2018
KX923803	*Entamoeba dispar*	non-human primate	NA	China	813	23-SEP-2017
KX923804	*Entamoeba dispar*	non-human primate	NA	China	843	23-SEP-2017
MF421530	*Entamoeba dispar*	*Homo sapiens*	stool	Iraq	353	13-AUG-2017
KX357142	*Entamoeba dispar*	*Homo sapiens*	pus aspirate	India	517	04-JUL-2016
KU320612	*Entamoeba dispar*	*Pan troglodytes schweinfurthii*	NA	NA	506	13-MAR-2016
KT825975	*Entamoeba dispar*	NA	NA	NA	573	26-OCT-2015
KT825976	*Entamoeba dispar*	NA	NA	NA	573	26-OCT-2015

KT825977	*Entamoeba dispar*	NA	NA	NA	573	26-OCT-2015
KT825978	*Entamoeba dispar*	NA	NA	NA	573	26-OCT-2015
KT825979	*Entamoeba dispar*	NA	NA	NA	573	26-OCT-2015
KT825980	*Entamoeba dispar*	NA	NA	NA	573	26-OCT-2015
KT825981	*Entamoeba dispar*	NA	NA	NA	573	26-OCT-2015
KT825982	*Entamoeba dispar*	NA	NA	NA	573	26-OCT-2015
KT825983	*Entamoeba dispar*	NA	NA	NA	573	26-OCT-2015
KP722596	*Entamoeba dispar*	*Homo sapiens*	stool	Iraq	1727	01-AUG-2015
KP722597	*Entamoeba dispar*	*Homo sapiens*	stool	Iraq	1714	01-AUG-2015
KP722598	*Entamoeba dispar*	*Homo sapiens*	stool	Iraq	1746	01-AUG-2015
KP722599	*Entamoeba dispar*	*Homo sapiens*	stool	Iraq	1728	01-AUG-2015
KP722600	*Entamoeba dispar*	*Homo sapiens*	stool	Iraq	1792	01-AUG-2015
KJ870214	*Entamoeba dispar*	NA	NA	Cameroon	536	17-JUN-2015
KJ870215	*Entamoeba dispar*	NA	NA	Cameroon	532	17-JUN-2015
KJ870216	*Entamoeba dispar*	NA	NA	Cameroon	535	17-JUN-2015
KJ870217	*Entamoeba dispar*	NA	NA	Cameroon	530	17-JUN-2015
KJ870218	*Entamoeba dispar*	NA	NA	Cameroon	534	17-JUN-2015
KJ870219	*Entamoeba dispar*	NA	NA	Cameroon	533	17-JUN-2015
KJ870220	*Entamoeba dispar*	NA	NA	Cameroon	536	17-JUN-2015
KJ870221	*Entamoeba dispar*	NA	NA	Cameroon	535	17-JUN-2015
KJ870222	*Entamoeba dispar*	NA	NA	Cameroon	531	17-JUN-2015
KJ870223	*Entamoeba dispar*	NA	NA	Cameroon	532	17-JUN-2015
KJ870224	*Entamoeba dispar*	NA	NA	Cameroon	535	17-JUN-2015
KJ870225	*Entamoeba dispar*	NA	NA	Cameroon	534	17-JUN-2015
KJ870226	*Entamoeba dispar*	NA	NA	Cameroon	532	17-JUN-2015
KJ870227	*Entamoeba dispar*	NA	NA	Cameroon	532	17-JUN-2015
KF646236	*Entamoeba dispar*	NA	NA	NA	445	30-MAR-2015
KF646237	*Entamoeba dispar*	NA	NA	NA	433	30-MAR-2015
KF646238	*Entamoeba dispar*	NA	NA	NA	377	30-MAR-2015
KJ719489	*Entamoeba dispar*	NA	NA	NA	573	13-JUL-2014
KJ188439	*Entamoeba dispar*	*Homo sapiens*	stool	India	239	15-APR-2014

HQ153408	*Entamoeba dispar*	NA	stool	Pakistan	170	25-AUG-2010
EF421340	*Entamoeba dispar*	NA	NA	NA	739	31-DEC-2007
EF421359	*Entamoeba dispar*	NA	NA	NA	777	31-DEC-2007
EF421360	*Entamoeba dispar*	NA	NA	NA	1522	31-DEC-2007
AB282661	*Entamoeba dispar*	NA	NA	NA	2426	25-APR-2007
EF204917	*Entamoeba dispar*	NA	NA	NA	366	31-JAN-2007
AY842965	*Entamoeba dispar*	NA	NA	NA	428	02-DEC-2005
AY842966	*Entamoeba dispar*	NA	NA	NA	318	02-DEC-2005
AY842968	*Entamoeba dispar*	NA	NA	NA	499	02-DEC-2005
Y12251	*Entamoeba dispar*	NA	NA	NA	423	04-APR-1997
AF031465	*Entamoeba dispar*	NA	NA	NA	313	02-JAN-1999
Z93402	*Entamoeba dispar*	NA	NA	NA	423	01-APR-1997
Z49256	*Entamoeba dispar*	NA	NA	NA	1949	01-MAR-1996

Table 11. *Entamoeba gingivalis*-specific rDNA sequences currently available in the NCBI nucleotide database (January 2023) listed according to date of deposition.*

NCBI Acc. no.	Organism	Subtype (ST)	Host	Isolation source	Country	Sequence length	Date of deposition
OP161459	*Entamoeba gingivalis*	ST1	*Homo sapiens*	oral cavity	Austria	436	01-JAN-2023
OP161460	*Entamoeba gingivalis*	ST1	*Homo sapiens*	oral cavity	Austria	434	01-JAN-2023
OP161461	*Entamoeba gingivalis*	ST1	*Homo sapiens*	oral cavity	Austria	434	01-JAN-2023
OP161462	*Entamoeba gingivalis*	ST1	*Homo sapiens*	oral cavity	Austria	434	01-JAN-2023
OP161463	*Entamoeba gingivalis*	ST1	*Homo sapiens*	oral cavity	Austria	434	01-JAN-2023
OP161464	*Entamoeba gingivalis*	ST1	*Homo sapiens*	oral cavity	Austria	434	01-JAN-2023
OP161465	*Entamoeba gingivalis*	ST1	*Homo sapiens*	oral cavity	Austria	434	01-JAN-2023
OP161466	*Entamoeba gingivalis*	ST1	*Homo sapiens*	oral cavity	Austria	434	01-JAN-2023
OP161467	*Entamoeba gingivalis*	ST1	*Homo sapiens*	oral cavity	Austria	434	01-JAN-2023
OP161468	*Entamoeba gingivalis*	ST1	*Homo sapiens*	oral cavity	Austria	438	01-JAN-2023
OP161469	*Entamoeba gingivalis*	ST3	*Homo sapiens*	oral cavity	Austria	457	01-JAN-2023
OP161470	*Entamoeba gingivalis*	ST3	*Homo sapiens*	oral cavity	Austria	457	01-JAN-2023
OP161471	*Entamoeba gingivalis*	ST3	*Homo sapiens*	oral cavity	Austria	457	01-JAN-2023
OP161472	*Entamoeba gingivalis*	ST2	*Homo sapiens*	oral cavity	Austria	456	01-JAN-2023
OP161473	*Entamoeba gingivalis*	ST2	*Homo sapiens*	oral cavity	Austria	453	01-JAN-2023
OP161474	*Entamoeba gingivalis*	ST2	*Homo sapiens*	oral cavity	Austria	453	01-JAN-2023
OP161475	*Entamoeba gingivalis*	ST2	*Homo sapiens*	oral cavity	Austria	451	01-JAN-2023
OP161476	*Entamoeba gingivalis*	ST2	*Homo sapiens*	oral cavity	Austria	454	01-JAN-2023
OP456213	*Entamoeba gingivalis*	ST1	*Homo sapiens*	oral cavity	Turkey	333	21-SEP-2022
OP456215	*Entamoeba gingivalis*	ST2	*Homo sapiens*	oral cavity	Turkey	355	21-SEP-2022
OP456304	*Entamoeba gingivalis*	ST1	*Homo sapiens*	oral cavity	Turkey	348	21-SEP-2022
OP422447	*Entamoeba gingivalis*	ST2	*Homo sapiens*	oral cavity	Turkey	348	18-SEP-2022
MZ388559	*Entamoeba gingivalis*	ST1	*Homo sapiens*	dental caries	Iraq	463	19-JUN-2021
MZ388560	*Entamoeba gingivalis*	ST1	*Homo sapiens*	dental caries	Iraq	454	19-JUN-2021
MZ388561	*Entamoeba gingivalis*	ST1	*Homo sapiens*	dental caries	Iraq	460	19-JUN-2021
MZ388562	*Entamoeba gingivalis*	ST1	*Homo sapiens*	dental caries	Iraq	451	19-JUN-2021
MZ388563	*Entamoeba gingivalis*	ST1	*Homo sapiens*	dental caries	Iraq	462	19-JUN-2021
MZ388564	*Entamoeba gingivalis*	ST1	*Homo sapiens*	dental caries	Iraq	460	19-JUN-2021
MW676260	*Entamoeba gingivalis*	ST1	Homo sapiens	mouthwash sample	Tanzania	370	07-MAR-2021

KX061778	*Entamoeba gingivalis*	ST1	*Homo sapiens*	dental plaque from HIV negative patient with gingivitis	Brazil	655	01-AUG-2019
KX061779	*Entamoeba gingivalis*	ST1	*Homo sapiens*	dental plaque from HIV negative patient with gingivitis	Brazil	660	01-AUG-2019
MG601094	*Entamoeba gingivalis*	ST1	*Homo sapiens*	female genital tract	USA	450	25-JUL-2018
KX027286	*Entamoeba gingivalis*	ST2	*Homo sapiens*	periodontal scraping from oral cavity	Mexico	369	31-JUL-2017
KX027287	*Entamoeba gingivalis*	ST2	*Homo sapiens*	periodontal scraping from oral cavity	Mexico	368	31-JUL-2017
KX027288	*Entamoeba gingivalis*	ST2	*Homo sapiens*	periodontal scraping from oral cavity	Mexico	369	31-JUL-2017
KX027289	*Entamoeba gingivalis*	ST2	*Homo sapiens*	periodontal scraping from oral cavity	Mexico	489	31-JUL-2017
KX027290	*Entamoeba gingivalis*	ST1	*Homo sapiens*	periodontal scraping from oral cavity	Mexico	520	31-JUL-2017
KX027291	*Entamoeba gingivalis*	ST1	*Homo sapiens*	periodontal scraping from oral cavity	Mexico	369	31-JUL-2017
KX027292	*Entamoeba gingivalis*	ST1	*Homo sapiens*	periodontal scraping from oral cavity	Mexico	368	31-JUL-2017
KX027293	*Entamoeba gingivalis*	ST1	*Homo sapiens*	periodontal scraping from oral cavity	Mexico	500	31-JUL-2017
KX027294	*Entamoeba gingivalis*	ST2	*Homo sapiens*	periodontal scraping from oral cavity	Mexico	2336	31-JUL-2017
KX027295	*Entamoeba gingivalis*	ST2	*Homo sapiens*	periodontal scraping from oral cavity	Mexico	2186	31-JUL-2017
KX027296	*Entamoeba gingivalis*	ST2	*Homo sapiens*	periodontal scraping from oral cavity	Mexico	2339	31-JUL-2017
KX027297	*Entamoeba gingivalis*	ST1	*Homo sapiens*	periodontal scraping from oral cavity	Mexico	2316	31-JUL-2017
KX027298	*Entamoeba gingivalis*	ST1	*Homo sapiens*	periodontal scraping from oral cavity	Mexico	2287	31-JUL-2017
KF250433	*Entamoeba gingivalis*	ST1	*Homo sapiens*	scraped saliva of human immunocompetent	Brazil	623	11-JAN-2014
KF250434	*Entamoeba gingivalis*	ST1	*Homo sapiens*	scraped saliva of human immunocompetent	Brazil	637	11-JAN-2014
KF250435	*Entamoeba gingivalis*	ST1	*Homo sapiens*	scraped saliva of HIV infected patient	Brazil	625	11-JAN-2014
KF250436	*Entamoeba gingivalis*	ST1	*Homo sapiens*	scraped saliva of HIV infected patient	Brazil	631	11-JAN-2014
D28490	*Entamoeba gingivalis*	ST1	*Homo sapiens*	subgingival space of adult with periodontal disease	NA	1918	25-MAR-1995

***** The sequence KU886548 is listed as 'Uncultured *Terfezia*' in GenBank; however, the sequence is *E. gingivalis* ST2 (not included in Table 11).

Table 12. *Entamoeba hartmanni*-specific SSU rDNA sequences currently available in the NCBI nucleotide database (January 2023) listed according to date of deposition. A few sequences in the database that are in fact *E. hartmanni* (Figure 10) are listed as '*Entamoeba* sp' (*e.g.,* KF515250, KF515240 and MG25651) or 'Uncultured *Entamoeba* clone' (*e.g.,* JX131943) and not included in the table. Information on subtype was identified where possible by aligning each sequence with the following near-complete sequences: AF149907 (ST1), KX618191 (ST1), FR686379 (ST1), FR686374 (ST1), FR686377 (ST2), FR686378 (ST3) and FR686376 (ST3). Quite a few of the sequences reflected the 5'-end of the SSU rRNA gene, and the subtype status of most of these remained unresolved; for a few of these, preliminary attempts were made to classify them.

NCBI Acc.no.	Organism	Subtype (ST)	Host	Isolation source	Country	Sequence length	Date of deposition
OP688358	*Entamoeba hartmanni*	Unresolved (similar to FR686369)	Domestic dogs	faeces	NA	529	25-OCT-2022
OP688359	*Entamoeba hartmanni*	ST1	Domestic dogs	faeces	NA	413	25-OCT-2022
OP688360	*Entamoeba hartmanni*	ST1	Domestic dogs	faeces	NA	413	25-OCT-2022
OP688361	*Entamoeba hartmanni*	ST1	Domestic dogs	faeces	NA	413	25-OCT-2022
OP688362	*Entamoeba hartmanni*	ST1	Domestic dogs	faeces	NA	413	25-OCT-2022
OP565047	*Entamoeba hartmanni*	Unresolved (similar to FR686369)	*Homo sapiens*	stool	NA	529	08-OCT-2022
OP565048	*Entamoeba hartmanni*	ST1	*Homo sapiens*	stool	NA	413	08-OCT-2022
OP565049	*Entamoeba hartmanni*	ST1	*Homo sapiens*	stool	NA	413	08-OCT-2022
OP565050	*Entamoeba hartmanni*	ST1	*Homo sapiens*	stool	NA	413	08-OCT-2022
OP565051	*Entamoeba hartmanni*	ST1	*Homo sapiens*	stool	NA	413	08-OCT-2022
ON974211	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	NA	413	18-JUL-2022
ON974212	*Entamoeba hartmanni*	ST3	*Homo sapiens*	faeces	NA	413	18-JUL-2022
ON974213	*Entamoeba hartmanni*	ST2	*Homo sapiens*	faeces	NA	412	18-JUL-2022
ON974214	*Entamoeba hartmanni*	ST2	*Homo sapiens*	faeces	NA	412	18-JUL-2022
ON974215	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	NA	413	18-JUL-2022
ON974216	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	NA	413	18-JUL-2022
ON974217	*Entamoeba hartmanni*	ST3	*Homo sapiens*	faeces	NA	413	18-JUL-2022
ON974218	*Entamoeba hartmanni*	ST3	*Homo sapiens*	faeces	NA	413	18-JUL-2022
ON974219	*Entamoeba hartmanni*	ST3	*Homo sapiens*	faeces	NA	413	18-JUL-2022
ON974220	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	NA	413	18-JUL-2022
ON974221	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	NA	413	18-JUL-2022
ON974222	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	NA	413	18-JUL-2022
ON974223	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	NA	413	18-JUL-2022
ON974224	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	NA	413	18-JUL-2022

ON974225	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	NA	413	18-JUL-2022
ON974226	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	NA	413	18-JUL-2022
ON974227	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	NA	413	18-JUL-2022
ON974228	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	NA	413	18-JUL-2022
ON974229	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	NA	413	18-JUL-2022
ON974230	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	NA	413	18-JUL-2022
ON974231	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	NA	413	18-JUL-2022
ON974232	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	NA	413	18-JUL-2022
ON974233	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	NA	413	18-JUL-2022
MT703882	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	Argentina	552	30-DEC-2021
MW026785	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	Brazil	487	30-MAR-2021
MW026786	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	Brazil	487	30-MAR-2021
MW026787	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	Brazil	487	30-MAR-2021
MW026788	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	Brazil	487	30-MAR-2021
MW026789	*Entamoeba hartmanni*	unresolved	*Homo sapiens*	faeces	Brazil	486	30-MAR-2021
MW026790	*Entamoeba hartmanni*	unresolved	*Homo sapiens*	faeces	Brazil	486	30-MAR-2021
MW026791	*Entamoeba hartmanni*	unresolved	*Homo sapiens*	faeces	Brazil	483	30-MAR-2021
MW026792	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	Brazil	488	30-MAR-2021
MK541027	*Entamoeba hartmanni*	ST1	*Homo sapiens*	faeces	Mexico	467	03-APR-2019
MH133212	*Entamoeba hartmanni*	ST1	rhesus macaques	faecal samples	China	590	01-APR-2019
MH623056	*Entamoeba hartmanni*	ST1	NA	waste water	Australia	108	24-MAR-2019
MH620470	*Entamoeba hartmanni*	Unresolved (only~94% similarity to closest match)	*Macaca cyclopis*	NA	Taiwan?	398	20-JUL-2018
MH620404	*Entamoeba hartmanni*	Unresolved (only~93% similarity to closest match)	*Macaca cyclopis*	NA	Taiwan?	279	19-JUL-2018
MF631986	*Entamoeba hartmanni*	ST1	NA	xenic culture	NA	178	28-MAR-2018
MG925065	*Entamoeba hartmanni*	ST1	*Homo sapiens*	stool	Tunisia	283	14-FEB-2018
MG925066	*Entamoeba hartmanni*	unresolved	*Homo sapiens*	stool	Tunisia	288	14-FEB-2018
MG925067	*Entamoeba hartmanni*	ST1	*Homo sapiens*	stool	Tunisia	283	14-FEB-2018
MG925068	*Entamoeba hartmanni*	ST1	*Homo sapiens*	stool	Tunisia	287	14-FEB-2018

MF421531	*Entamoeba hartmanni*	ST1	*Homo sapiens*	stool	Iraq	429	13-AUG-2017
MF471207	*Entamoeba hartmanni*	unresolved	*Homo sapiens*	NA	South Africa	121	18-JUL-2017
MF471208	*Entamoeba hartmanni*	unresolved	*Homo sapiens*	NA	South Africa	121	18-JUL-2017
MF471209	*Entamoeba hartmanni*	unresolved	*Homo sapiens*	NA	South Africa	122	18-JUL-2017
MF471210	*Entamoeba hartmanni*	unresolved	*Homo sapiens*	NA	South Africa	121	18-JUL-2017
MF471211	*Entamoeba hartmanni*	ST1	*Homo sapiens*	NA	South Africa	122	18-JUL-2017
MF471212	*Entamoeba hartmanni*	unresolved	*Homo sapiens*	NA	South Africa	121	18-JUL-2017
MF471213	*Entamoeba hartmanni*	ST1	*Homo sapiens*	NA	South Africa	122	18-JUL-2017
MF471214	*Entamoeba hartmanni*	ST1	*Homo sapiens*	NA	South Africa	122	18-JUL-2017
MF471215	*Entamoeba hartmanni*	ST1	*Homo sapiens*	NA	South Africa	122	18-JUL-2017
MF471216	*Entamoeba hartmanni*	ST3	*Homo sapiens*	NA	South Africa	121	18-JUL-2017
MF471217	*Entamoeba hartmanni*	unresolved	*Homo sapiens*	NA	South Africa	121	18-JUL-2017
KX618191	*Entamoeba hartmanni*	ST1	*Homo sapiens*	stool	Singapore	1910	04-APR-2017
KU320609	*Entamoeba hartmanni*	unresolved	*Pan troglodytes schweinfurthii*	NA	NA	544	13-MAR-2016
FR686366	*Entamoeba hartmanni*	ST1	*Lagothrix lagotricha*	host faeces	United Kingdom	530	18-OCT-2010
FR686367	*Entamoeba hartmanni*	ST1	*Lagothrix lagotricha*	host faeces	United Kingdom	530	18-OCT-2010
FR686368	*Entamoeba hartmanni*	ST1	*Lagothrix lagotricha*	host faeces	United Kingdom	530	18-OCT-2010
FR686369	*Entamoeba hartmanni*	unresolved	*Macaca sylvanus*	host faeces	United Kingdom	529	18-OCT-2010
FR686370	*Entamoeba hartmanni*	unresolved	*Pongo pygmaeus*	host faeces	United Kingdom	530	18-OCT-2010
FR686371	*Entamoeba hartmanni*	unresolved	*Papio* sp.	host faeces	United Kingdom	529	18-OCT-2010
FR686372	*Entamoeba hartmanni*	ST1	*Macaca sylvanus*	host faeces	United Kingdom	530	18-OCT-2010
FR686373	*Entamoeba hartmanni*	ST1	*Erythrocebus patas*	host faeces	United Kingdom	530	18-OCT-2010
FR686374	*Entamoeba hartmanni*	ST1	*Homo sapiens*	host faeces	NA	1900	18-OCT-2010
FR686375	*Entamoeba hartmanni*	ST1	*Homo sapiens*	host faeces	NA	1491	18-OCT-2010
FR686376	*Entamoeba hartmanni*	ST3	*Homo sapiens*	host faeces	NA	1897	18-OCT-2010
FR686377	*Entamoeba hartmanni*	ST2	*Homo sapiens*	host faeces	NA	1899	18-OCT-2010
FR686378	*Entamoeba hartmanni*	ST3	*Homo sapiens*	host faeces	NA	1898	18-OCT-2010
FR686379	*Entamoeba hartmanni*	ST1	*Homo sapiens*	host faeces	NA	1900	18-OCT-2010
FR686380	*Entamoeba hartmanni*	ST1	*Homo sapiens*	host faeces	NA	1491	18-OCT-2010
FR686381	*Entamoeba hartmanni*	ST3	*Homo sapiens*	host faeces	NA	1489	18-OCT-2010

FR686382	*Entamoeba hartmanni*	ST1	*Homo sapiens*	host faeces	NA	530	18-OCT-2010
AF149907	*Entamoeba hartmanni*	ST1	NA	NA	NA	1960	22-DEC-1999

Table 13. *Entamoeba moshkovskii*-specific SSU rDNA sequences currently available in the NCBI nucleotide database (January 2023) listed according to date of deposition.

NCBI Acc.no.	Organism	Host	Isolation source	Country	Sequence length	Date of deposition
OP537191	*Entamoeba moshkovskii*	*Homo sapiens*	stool	NA	733	02-OCT-2022
OP537192	*Entamoeba moshkovskii*	*Homo sapiens*	stool	NA	733	02-OCT-2022
OP537193	*Entamoeba moshkovskii*	*Homo sapiens*	stool	NA	733	02-OCT-2022
OP537194	*Entamoeba moshkovskii*	*Homo sapiens*	stool	NA	733	02-OCT-2022
OP537195	*Entamoeba moshkovskii*	*Homo sapiens*	stool	NA	733	02-OCT-2022
OP537196	*Entamoeba moshkovskii*	*Homo sapiens*	stool	NA	733	02-OCT-2022
OP537197	*Entamoeba moshkovskii*	*Homo sapiens*	stool	NA	733	02-OCT-2022
OP537198	*Entamoeba moshkovskii*	*Homo sapiens*	stool	NA	733	02-OCT-2022
OP537199	*Entamoeba moshkovskii*	*Homo sapiens*	stool	NA	733	02-OCT-2022
OP537200	*Entamoeba moshkovskii*	*Homo sapiens*	stool	NA	733	02-OCT-2022
OP529844	*Entamoeba moshkovskii*	*Homo sapiens*	stool	NA	539	01-OCT-2022
OP529845	*Entamoeba moshkovskii*	*Homo sapiens*	stool	NA	536	01-OCT-2022
OP529846	*Entamoeba moshkovskii*	*Homo sapiens*	stool	NA	533	01-OCT-2022
OP529847	*Entamoeba moshkovskii*	*Homo sapiens*	stool	NA	535	01-OCT-2022
OP529848	*Entamoeba moshkovskii*	*Homo sapiens*	stool	NA	536	01-OCT-2022
OP452928	*Entamoeba moshkovskii*	*Homo sapiens*	stool	NA	560	21-SEP-2022
OP452929	*Entamoeba moshkovskii*	*Homo sapiens*	stool	NA	550	21-SEP-2022
OP452930	*Entamoeba moshkovskii*	*Homo sapiens*	stool	NA	1724	21-SEP-2022
OP452931	*Entamoeba moshkovskii*	*Homo sapiens*	stool	NA	1743	21-SEP-2022
OP452932	*Entamoeba moshkovskii*	*Homo sapiens*	stool	NA	1741	21-SEP-2022
OP453350	*Entamoeba moshkovskii*	Human and Domestic dogs	stool	NA	733	21-SEP-2022
OP453351	*Entamoeba moshkovskii*	Human and Domestic dogs	stool	NA	733	21-SEP-2022
OP453352	*Entamoeba moshkovskii*	Human and Domestic dogs	stool	NA	733	21-SEP-2022
OP453353	*Entamoeba moshkovskii*	Human and Domestic dogs	stool	NA	733	21-SEP-2022
OP453354	*Entamoeba moshkovskii*	Human and Domestic dogs	stool	NA	733	21-SEP-2022
ON965383	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965384	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965385	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022

ON965386	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965387	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965388	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965389	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965390	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965391	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965392	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965393	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965394	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965395	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965396	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965397	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965398	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965399	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965400	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965401	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965402	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965403	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965404	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965405	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965406	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965407	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965408	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965409	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965410	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965411	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965412	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965413	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965414	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965415	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022

ON965416	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965417	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965418	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965419	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965420	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965421	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965422	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965423	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965424	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965425	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965426	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965427	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965428	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965429	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965430	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965431	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965432	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965433	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965434	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965435	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965436	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965437	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965438	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965439	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965440	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965441	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965442	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965443	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965444	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965445	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022

ON965446	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965447	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965448	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965449	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON965450	*Entamoeba moshkovskii*	*Homo sapiens*	faecal samples	India	733	17-JUL-2022
ON318868	*Entamoeba moshkovskii*	NA	faeces	Malaysia	553	27-APR-2022
ON318869	*Entamoeba moshkovskii*	NA	faeces	Malaysia	553	27-APR-2022
ON318870	*Entamoeba moshkovskii*	NA	faeces	Malaysia	553	27-APR-2022
ON318871	*Entamoeba moshkovskii*	NA	faeces	Malaysia	553	27-APR-2022
ON318872	*Entamoeba moshkovskii*	NA	faeces	Malaysia	553	27-APR-2022
OM791622	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	1145	28-FEB-2022
OM791623	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	1145	28-FEB-2022
OM791624	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	1145	28-FEB-2022
OM791625	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	1145	28-FEB-2022
OM791626	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	1145	28-FEB-2022
OM791627	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	1145	28-FEB-2022
OM791628	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	1145	28-FEB-2022
OM791629	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	1145	28-FEB-2022
OM791630	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	1145	28-FEB-2022
OM791631	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	1145	28-FEB-2022
MZ913026	*Entamoeba moshkovskii*	Human, Child	stool	Iraq	549	30-AUG-2021
MZ913027	*Entamoeba moshkovskii*	Human, Child	stool	Iraq	549	30-AUG-2021
MZ357989	*Entamoeba moshkovskii*	pig	faeces	India	222	13-JUN-2021
MZ357990	*Entamoeba moshkovskii*	pig	faeces	India	222	13-JUN-2021
MZ357991	*Entamoeba moshkovskii*	pig	faeces	India	222	13-JUN-2021
MZ357992	*Entamoeba moshkovskii*	pig	faeces	India	222	13-JUN-2021
MZ357993	*Entamoeba moshkovskii*	pig	faeces	India	222	13-JUN-2021
MZ357994	*Entamoeba moshkovskii*	pig	faeces	India	222	13-JUN-2021
MZ357995	*Entamoeba moshkovskii*	pig	faeces	India	222	13-JUN-2021
MZ357996	*Entamoeba moshkovskii*	pig	faeces	India	222	13-JUN-2021

MZ357997	*Entamoeba moshkovskii*	pig	faeces	India	222	13-JUN-2021
MZ357998	*Entamoeba moshkovskii*	pig	faeces	India	222	13-JUN-2021
MZ357999	*Entamoeba moshkovskii*	pig	faeces	India	222	13-JUN-2021
MW926950	*Entamoeba moshkovskii*	pig	faeces	India	222	20-APR-2021
MW309341	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309342	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309343	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309344	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309345	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309346	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309347	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309348	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309349	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309350	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309351	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309352	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309353	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309354	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309355	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309356	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309357	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309358	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309359	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309360	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309361	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309362	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309363	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309364	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309365	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309366	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020

MW309367	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309368	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309369	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309370	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309371	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309372	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309373	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309374	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309375	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309376	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309377	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309378	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309379	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309380	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309381	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MW309382	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	07-DEC-2020
MN749976	*Entamoeba moshkovskii*	NA	NA	NA	860	01-DEC-2020
MW165340	*Entamoeba moshkovskii*	*Homo sapiens*	stool	Iraq	528	28-OCT-2020
MT350103	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	22-APR-2020
MT350104	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	22-APR-2020
MT350105	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	22-APR-2020
MT350106	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	22-APR-2020
MT350107	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	22-APR-2020
MT350108	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	22-APR-2020
MT350109	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	22-APR-2020
MT350110	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	223	22-APR-2020
MT350111	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	22-APR-2020
MT350112	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	22-APR-2020
MT350113	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	22-APR-2020
MT350114	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	22-APR-2020

MT350115	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	22-APR-2020
MT350116	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	22-APR-2020
MT350117	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	222	22-APR-2020
MT250838	*Entamoeba moshkovskii*	*Homo sapiens*	stool	Iraq	561	31-MAR-2020
MN498050	*Entamoeba moshkovskii*	environmental	wastewater	NA	605	15-JAN-2020
MN498051	*Entamoeba moshkovskii*	environmental	wastewater	NA	606	15-JAN-2020
MK142734	*Entamoeba moshkovskii*	*Homo sapiens*	stool	Kenya	495	04-DEC-2019
MK142735	*Entamoeba moshkovskii*	*Homo sapiens*	stool	Kenya	806	04-DEC-2019
MK142736	*Entamoeba moshkovskii*	*Homo sapiens*	stool	Kenya	584	04-DEC-2019
MK142737	*Entamoeba moshkovskii*	*Homo sapiens*	stool	Kenya	601	04-DEC-2019
MN536488	*Entamoeba moshkovskii*	snake	NA	Czech Republic	1938	10-OCT-2019
MN536489	*Entamoeba moshkovskii*	NA	NA	NA	1938	10-OCT-2019
MN536490	*Entamoeba moshkovskii*	environmental	costal marine sediment	USA	1937	10-OCT-2019
MN536491	*Entamoeba moshkovskii*	environmental	freshwater stream sediment	USA	1938	10-OCT-2019
MN536492	*Entamoeba moshkovskii*	*Periplaneta americana*	hind gut	USA	1940	10-OCT-2019
MN536493	*Entamoeba moshkovskii*	environmental	freshwater pond sediment	USA	1940	10-OCT-2019
MN536494	*Entamoeba moshkovskii*	environmental	brackish water estuary sediment	USA	1938	10-OCT-2019
MN536495	*Entamoeba moshkovskii*	*Cotinis nitida*	hind gut	USA	1565	10-OCT-2019
MN536496	*Entamoeba moshkovskii*	environmental	costal marine sediment	USA	1565	10-OCT-2019
MN536497	*Entamoeba moshkovskii*	environmental	freshwater sediment	USA	1564	10-OCT-2019
MN536498	*Entamoeba moshkovskii*	environmental	freshwater sediment	USA	1565	10-OCT-2019
MN536499	*Entamoeba moshkovskii*	environmental	freshwater sediment	USA	1564	10-OCT-2019
MN536500	*Entamoeba moshkovskii*	environmental	freshwater sediment	USA	1564	10-OCT-2019
MN536501	*Entamoeba moshkovskii*	environmental	freshwater sediment	USA	1563	10-OCT-2019
MN536502	*Entamoeba moshkovskii*	environmental	freshwater sediment	USA	1564	10-OCT-2019
MN535795	*Entamoeba moshkovskii*	*Periplaneta americana*	DNA from hind gut	NA	1527	09-OCT-2019
MN535796	*Entamoeba moshkovskii*	*Periplaneta americana*	DNA from hind gut	NA	1527	09-OCT-2019
MN496101	*Entamoeba moshkovskii*	Homo sapiens	stool	India	222	28-SEP-2019
MN496102	*Entamoeba moshkovskii*	Homo sapiens	stool	India	222	28-SEP-2019

MN496103	*Entamoeba moshkovskii*	Homo sapiens	stool	India	222	28-SEP-2019
MN496104	*Entamoeba moshkovskii*	Homo sapiens	stool	India	222	28-SEP-2019
MN496105	*Entamoeba moshkovskii*	Homo sapiens	stool	India	222	28-SEP-2019
MN496106	*Entamoeba moshkovskii*	Homo sapiens	stool	India	222	28-SEP-2019
MN496107	*Entamoeba moshkovskii*	Homo sapiens	stool	India	222	28-SEP-2019
MN496108	*Entamoeba moshkovskii*	Homo sapiens	stool	India	222	28-SEP-2019
MN496109	*Entamoeba moshkovskii*	Homo sapiens	stool	India	222	28-SEP-2019
MN496110	*Entamoeba moshkovskii*	Homo sapiens	stool	India	222	28-SEP-2019
MN496111	*Entamoeba moshkovskii*	Homo sapiens	stool	India	222	28-SEP-2019
MN496112	*Entamoeba moshkovskii*	Homo sapiens	stool	India	222	28-SEP-2019
MN496113	*Entamoeba moshkovskii*	Homo sapiens	stool	India	222	28-SEP-2019
MN496114	*Entamoeba moshkovskii*	Homo sapiens	stool	India	222	28-SEP-2019
MN496115	*Entamoeba moshkovskii*	Homo sapiens	stool	India	222	28-SEP-2019
MN496116	*Entamoeba moshkovskii*	Homo sapiens	stool	India	222	28-SEP-2019
MN496117	*Entamoeba moshkovskii*	Homo sapiens	stool	India	222	28-SEP-2019
MN496118	*Entamoeba moshkovskii*	Homo sapiens	stool	India	222	28-SEP-2019
MN496119	*Entamoeba moshkovskii*	Homo sapiens	stool	India	222	28-SEP-2019
MH623061	*Entamoeba moshkovskii*	environmental	waste water	Australia	105	24-MAR-2019
MF631987	*Entamoeba moshkovskii*	environmental	soil; xenic culture	NA	255	28-MAR-2018
MF631988	*Entamoeba moshkovskii*	NA	axenic culture	NA	203	28-MAR-2018
KY823428	*Entamoeba moshkovskii*	NA	NA	NA	494	26-FEB-2018
KT825984	*Entamoeba moshkovskii*	NA	NA	NA	509	26-OCT-2015
KT825985	*Entamoeba moshkovskii*	NA	NA	NA	509	26-OCT-2015
KT825986	*Entamoeba moshkovskii*	NA	NA	NA	509	26-OCT-2015
KT825987	*Entamoeba moshkovskii*	NA	NA	NA	509	26-OCT-2015
KT825988	*Entamoeba moshkovskii*	NA	NA	NA	509	26-OCT-2015
KT825989	*Entamoeba moshkovskii*	NA	NA	NA	509	26-OCT-2015
KT825990	*Entamoeba moshkovskii*	NA	NA	NA	509	26-OCT-2015
KT825991	*Entamoeba moshkovskii*	NA	NA	NA	509	26-OCT-2015
KT825992	*Entamoeba moshkovskii*	NA	NA	NA	509	26-OCT-2015

KT825993	*Entamoeba moshkovskii*	NA	NA	NA	509	26-OCT-2015
KP722601	*Entamoeba moshkovskii*	*Homo sapiens*	stool	Iraq	1741	01-AUG-2015
KP722602	*Entamoeba moshkovskii*	*Homo sapiens*	stool	Iraq	1781	01-AUG-2015
KP722603	*Entamoeba moshkovskii*	*Homo sapiens*	stool	Iraq	1743	01-AUG-2015
KP722604	*Entamoeba moshkovskii*	*Homo sapiens*	stool	Iraq	1724	01-AUG-2015
KP722605	*Entamoeba moshkovskii*	*Homo sapiens*	stool	Iraq	1743	01-AUG-2015
KJ870228	*Entamoeba moshkovskii*	NA	NA	Cameroon	539	17-JUN-2015
KJ870229	*Entamoeba moshkovskii*	NA	NA	Cameroon	536	17-JUN-2015
KJ870230	*Entamoeba moshkovskii*	NA	NA	Cameroon	533	17-JUN-2015
KJ870231	*Entamoeba moshkovskii*	NA	NA	Cameroon	535	17-JUN-2015
KJ870232	*Entamoeba moshkovskii*	NA	NA	Cameroon	536	17-JUN-2015
KJ870233	*Entamoeba moshkovskii*	NA	NA	Cameroon	537	17-JUN-2015
KJ776599	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	509	29-JUL-2014
KJ776600	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	498	29-JUL-2014
KJ719493	*Entamoeba moshkovskii*	NA	NA	NA	509	13-JUL-2014
KJ719494	*Entamoeba moshkovskii*	NA	NA	NA	498	13-JUL-2014
JQ406865	*Entamoeba moshkovskii*	NA	NA	NA	462	28-APR-2012
JQ406866	*Entamoeba moshkovskii*	NA	NA	NA	462	28-APR-2012
JQ406867	*Entamoeba moshkovskii*	NA	NA	NA	462	28-APR-2012
AB520687	*Entamoeba moshkovskii*	*Homo sapiens*	stool	Iran	579	08-SEP-2009
EF421285	*Entamoeba moshkovskii*	NA	NA	NA	559	31-DEC-2007
EF421303	*Entamoeba moshkovskii*	NA	NA	NA	676	31-DEC-2007
EF421305	*Entamoeba moshkovskii*	NA	NA	NA	983	31-DEC-2007
EF421366	*Entamoeba moshkovskii*	NA	NA	NA	575	31-DEC-2007
EF682206	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	480	05-JUL-2007
EF682207	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	480	05-JUL-2007
EF682208	*Entamoeba moshkovskii*	*Homo sapiens*	stool	India	480	05-JUL-2007
EF204916	*Entamoeba moshkovskii*	NA	NA	NA	136	31-JAN-2007
AF149906	*Entamoeba moshkovskii*	NA	NA	NA	1944	22-DEC-1999
X89635	*Entamoeba moshkovskii*	NA	NA	NA	490	19-JUL-1995

Table 14. Host specificity of *Entamoeba polecki* subtypes 1 to 4 (updated version of the Supplementary Table included in the study by Stensvold et al., 2018 [53]).

E. polecki subtype (ST)	Host species	References/GenBank accession number
ST1	*Homo sapiens*	[257]
	Sus scrofa	[13, 30, 53, 116, 349-351]
	Rhea americana	[49]
	*Rhinopithecus bieti**	[346]
	Felis catus	OP753638 (Direct submission, no accompanying article)
ST2	*Homo sapiens*	[53, 257]
	Macaca mulata	[352]
	Macaca fascicularis	[49]
	Macaca fuscata fuscata	[266]
	Papio cynomolgi	[30]
	Leontopithecus chrysomelas	[353]
	Trachypithecus auratus auratus	[353]
ST3	*Homo sapiens*	[53, 257]
	Sus scrofa	[13, 53, 116, 351]
	Struthio camelus	[354]
	Rhea americana	[49]
ST4	*Homo sapiens*	[49, 53, 257, 355]
	Macaca nigra	[349]
Unresolved	"Dog"	OP919601 (Direct submission, no accompanying article)

* The sequence MW718195 was annotated as *E. polecki* ST1 by the authors [346]. Indeed, this taxon is the closest match by BLAST; however, there is only 97.30% match to the most similar sequence, so this host species (*Rhinopithecus bieti*) remains unconfirmed. The sequence is only 259 bp long, but could be considered a new conditional lineage.

Table 15. *Entamoeba polecki*-specific SSU rDNA sequences currently available in the NCBI nucleotide database (January 2023) listed according to date of deposition. In the second column, the entries for 'Organisms' were those originally entered in the NCBI database. *Entamoeba polecki* includes *E. polecki* subtypes 1 to 4, whereas *Entamoeba chattoni* and *Entamoeba struthionis* are redundant names for *E. polecki* ST2 and *E. polecki* ST3, respectively.

NCBI Acc.no.	Organism	Subtype	Host	Isolation source	Country	Sequence length	Date of deposition
OP919601	*Entamoeba polecki*	unresolved	dog	faeces	India	578	05-DEC-2022
OP753638	*Entamoeba polecki*	ST1	*Felis catus*	NA	India	578	08-NOV-2022
OP564991	*Entamoeba polecki*	ST3	*Homo sapiens*	stool	NA	560	08-OCT-2022
OP564992	*Entamoeba polecki*	ST2	*Homo sapiens*	stool	NA	554	08-OCT-2022
OP564993	*Entamoeba polecki*	ST3	*Homo sapiens*	stool	NA	550	08-OCT-2022
OP564994	*Entamoeba polecki*	ST2	*Homo sapiens*	stool	NA	554	08-OCT-2022
OP564995	*Entamoeba polecki*	ST4	*Homo sapiens*	stool	NA	554	08-OCT-2022
OP010911	*Entamoeba polecki*	ST3	pig	faeces	India	620	24-JUL-2022
MW718195	*Entamoeba polecki*	unresolved	*Rhinopithecus bieti*	NA	NA	259	17-AUG-2021
LC230016	*Entamoeba polecki*	ST1	*Sus scrofa domesticus*	faeces	Japan	1806	18-OCT-2019
LC230017	*Entamoeba polecki*	ST2	*Macaca fuscata fuscata*	faeces	Japan	1811	18-OCT-2019
LC230018	*Entamoeba polecki*	ST3	*Sus scrofa domesticus*	faeces	Japan	1811	18-OCT-2019
MK801424	*Entamoeba polecki*	ST1	*Sus scrofa domesticus*	faeces	Germany	1809	13-SEP-2019
MK801425	*Entamoeba polecki*	ST3	*Sus scrofa domesticus*	faeces	Germany	1801	13-SEP-2019
MK801426	*Entamoeba polecki*	ST1	*Sus scrofa domesticus*	faeces	Germany	1824	13-SEP-2019
MK801427	*Entamoeba polecki*	ST3	*Sus scrofa domesticus*	faeces	Germany	1604	13-SEP-2019

MK801429	*Entamoeba polecki*	ST1	*Sus scrofa domesticus*	faeces	Germany	1821	13-SEP-2019
MK801430	*Entamoeba polecki*	ST3	*Sus scrofa domesticus*	faeces	Germany	1713	13-SEP-2019
MK801433	*Entamoeba polecki*	ST3	*Sus scrofa domesticus*	faeces	Germany	1799	13-SEP-2019
MK801435	*Entamoeba polecki*	ST1	*Sus scrofa domesticus*	faeces	Germany	1832	13-SEP-2019
MK801439	*Entamoeba polecki*	ST1	*Sus scrofa domesticus*	faeces	Germany	1800	13-SEP-2019
MK801440	*Entamoeba polecki*	ST3	*Sus scrofa domesticus*	faeces	Germany	1724	13-SEP-2019
MK801442	*Entamoeba polecki*	ST1	*Sus scrofa domesticus*	faeces	Germany	1828	13-SEP-2019
MK801443	*Entamoeba polecki*	ST3	*Sus scrofa domesticus*	faeces	Germany	1714	13-SEP-2019
MK801446	*Entamoeba polecki*	ST1	*Sus scrofa domesticus*	faeces	Germany	1798	13-SEP-2019
MK801449	*Entamoeba polecki*	ST1	*Sus scrofa domesticus*	faeces	Germany	1834	13-SEP-2019
MK801450	*Entamoeba polecki*	ST1-ST3 chimaera	*Sus scrofa domesticus*	faeces	Germany	1805	13-SEP-2019
MK801452	*Entamoeba polecki*	ST1	*Sus scrofa domesticus*	faeces	Austria	1811	13-SEP-2019
MK801453	*Entamoeba polecki*	ST3	*Sus scrofa domesticus*	faeces	Austria	1727	13-SEP-2019
MK801455	*Entamoeba polecki*	ST1	*Sus scrofa domesticus*	faeces	Austria	1816	13-SEP-2019
MK801456	*Entamoeba polecki*	ST1	*Sus scrofa domesticus*	faeces	Austria	1829	13-SEP-2019
MK801457	*Entamoeba polecki*	ST1	*Sus scrofa domesticus*	faeces	Romania	1811	13-SEP-2019
MK801458	*Entamoeba polecki*	ST3	*Sus scrofa domesticus*	faeces	Romania	1738	13-SEP-2019
MK801460	*Entamoeba polecki*	ST1	*Sus scrofa domesticus*	faeces	Romania	1815	13-SEP-2019
MK801461	*Entamoeba polecki*	ST3	*Sus scrofa domesticus*	faeces	Romania	1722	13-SEP-2019

MK347347	*Entamoeba polecki*	ST3	swine	NA	China	187	04-AUG-2019
MK357717	*Entamoeba polecki*	ST1	swine	NA	China	201	24-JUL-2019
MH623051	*Entamoeba polecki*	ST4	NA	waste water	Australia	109	24-MAR-2019
MK559469	*Entamoeba polecki*	ST2	rhesus monkeys	faeces	NA	202	27-FEB-2019
MK559470	*Entamoeba polecki*	ST2	rhesus monkeys	faeces	NA	200	27-FEB-2019
MK559471	*Entamoeba polecki*	ST2	rhesus monkeys	faeces	NA	203	27-FEB-2019
MH011332	*Entamoeba polecki*	ST3	NA	faecal	China	805	26-FEB-2019
MH011333	*Entamoeba polecki*	ST1	NA	faecal	China	800	26-FEB-2019
MH348163	*Entamoeba polecki*	ST1	*Sus scrofa domestica*	host faeces	Argentina	202	31-DEC-2018
MH348164	*Entamoeba polecki*	ST1	*Sus scrofa domestica*	host faeces	Argentina	202	31-DEC-2018
MH348165	*Entamoeba polecki*	ST3	*Sus scrofa domestica*	host faeces	Argentina	204	31-DEC-2018
MH348166	*Entamoeba polecki*	ST3	*Sus scrofa domestica*	host faeces	Argentina	204	31-DEC-2018
MH348167	*Entamoeba polecki*	ST3	*Sus scrofa domestica*	host faeces	Argentina	204	31-DEC-2018
MH348168	*Entamoeba polecki*	ST3	*Sus scrofa domestica*	host faeces	Argentina	204	31-DEC-2018
MH348169	*Entamoeba polecki*	ST1	*Sus scrofa domestica*	host faeces	Argentina	202	31-DEC-2018
MH348170	*Entamoeba polecki*	ST1	*Sus scrofa domestica*	host faeces	Argentina	202	31-DEC-2018
MH348171	*Entamoeba polecki*	ST1	*Sus scrofa domestica*	host faeces	Argentina	202	31-DEC-2018
MH348172	*Entamoeba polecki*	ST3	*Sus scrofa domestica*	host faeces	Argentina	204	31-DEC-2018
MH348173	*Entamoeba polecki*	ST3	*Sus scrofa domestica*	host faeces	Argentina	204	31-DEC-2018

MH348174	*Entamoeba polecki*	ST1	*Sus scrofa domestica*	host faeces	Argentina	202	31-DEC-2018
MH348175	*Entamoeba polecki*	ST1	*Sus scrofa domestica*	host faeces	Argentina	202	31-DEC-2018
MG601093	*Entamoeba polecki*	ST4	*Homo sapiens*	Female genital tract	USA	395	25-JUL-2018
MG747649	*Entamoeba polecki*	ST4	NA	NA	NA	553	01-MAR-2018
MG747650	*Entamoeba polecki*	ST4	NA	NA	NA	540	01-MAR-2018
MG747651	*Entamoeba polecki*	ST4	NA	NA	NA	543	01-MAR-2018
MG747652	*Entamoeba polecki*	ST4	NA	NA	NA	550	01-MAR-2018
MG747653	*Entamoeba polecki*	ST3	NA	NA	NA	554	01-MAR-2018
MG747654	*Entamoeba polecki*	ST4	NA	NA	NA	589	01-MAR-2018
MG747655	*Entamoeba polecki*	ST4	NA	NA	NA	589	01-MAR-2018
MG747656	*Entamoeba polecki*	ST4	NA	NA	NA	548	01-MAR-2018
MG747657	*Entamoeba polecki*	ST2	NA	NA	NA	590	01-MAR-2018
MG747658	*Entamoeba polecki*	ST4	NA	NA	NA	549	01-MAR-2018
MG747659	*Entamoeba polecki*	ST4	NA	NA	NA	555	01-MAR-2018
MG747660	*Entamoeba polecki*	ST4	NA	NA	NA	589	01-MAR-2018
MG747661	*Entamoeba polecki*	ST4	NA	NA	NA	589	01-MAR-2018
MG747662	*Entamoeba polecki*	ST4	NA	NA	NA	589	01-MAR-2018
MG747663	*Entamoeba polecki*	ST4	NA	NA	NA	589	01-MAR-2018
MG747664	*Entamoeba polecki*	ST4	NA	NA	NA	554	01-MAR-2018

MG747665	*Entamoeba polecki*	ST2	NA	NA	NA	554	01-MAR-2018
MG747666	*Entamoeba polecki*	ST3	NA	NA	NA	550	01-MAR-2018
MG747667	*Entamoeba polecki*	ST2	NA	NA	NA	554	01-MAR-2018
MG747668	*Entamoeba polecki*	ST3	NA	NA	NA	560	01-MAR-2018
LC082304	*Entamoeba polecki*	ST4	*Macaca nigra*	NA	Indonesia	1817	10-MAR-2016
LC082305	*Entamoeba polecki*	ST1	*Sus scrofa domesticus*	NA	Indonesia	1628	10-MAR-2016
AB851498	*Entamoeba polecki*	ST4	*Homo sapiens*	HIV patient	Cameroon	324	23-FEB-2016
AB845670	*Entamoeba polecki*	Not *Entamoeba*	*Homo sapiens*	HIV patient	Cameroon	197	16-JAN-2016
AB845671	*Entamoeba polecki*	Not *Entamoeba*	*Homo sapiens*	HIV patient	Cameroon	171	16-JAN-2016
FR686357	*Entamoeba polecki*	ST4	*Homo sapiens*	host faeces	Netherlands	1848	18-OCT-2010
FR686383	*Entamoeba polecki*	ST1	*Homo sapiens*	host faeces	Denmark	1539	18-OCT-2010
FR686384	*Entamoeba polecki*	ST1	*Sus scrofa domesticus*	host faeces	Denmark	548	18-OCT-2010
FR686385	*Entamoeba polecki*	ST3	*Sus scrofa domesticus*	host faeces	Denmark	562	18-OCT-2010
FR686386	*Entamoeba polecki*	ST3	*Sus scrofa domesticus*	host faeces	NA	561	18-OCT-2010
FR686387	*Entamoeba polecki*	ST1	*Rhea americana*	purified cysts from faeces	Sweden	578	18-OCT-2010
FR686388	*Entamoeba polecki*	ST3	*Rhea americana*	purified cysts from faeces	Sweden	537	18-OCT-2010
FR686389	*Entamoeba polecki*	ST2	*Macaca fascicularis*	host faeces	NA	550	18-OCT-2010
FR686390	*Entamoeba polecki*	ST3	*Homo sapiens*	host faeces	Sweden	1197	18-OCT-2010
FR686391	*Entamoeba polecki*	ST3	*Homo sapiens*	host faeces	Nigeria	538	18-OCT-2010

FR686392	*Entamoeba polecki*	ST4	*Homo sapiens*	host faeces	Somalia	1812	18-OCT-2010
FR686393	*Entamoeba polecki*	ST4	*Homo sapiens*	host faeces	NA	1825	18-OCT-2010
FR686394	*Entamoeba polecki*	ST4	*Homo sapiens*	host faeces	Ethiopia	1812	18-OCT-2010
FR686395	*Entamoeba polecki*	ST4	*Homo sapiens*	host faeces	Sudan	1806	18-OCT-2010
FR686396	*Entamoeba polecki*	ST4	*Homo sapiens*	host faeces	NA	549	18-OCT-2010
FR686397	*Entamoeba polecki*	ST4	*Homo sapiens*	host faeces	Viet Nam	1810	18-OCT-2010
FR686398	*Entamoeba polecki*	ST4	*Homo sapiens*	host faeces	Kenya	1826	18-OCT-2010
FR686399	*Entamoeba polecki*	ST4	*Homo sapiens*	host faeces	Iraq	546	18-OCT-2010
FR686400	*Entamoeba polecki*	ST4	*Homo sapiens*	host faeces	NA	1823	18-OCT-2010
LC018995	*Entamoeba polecki*	ST1	*Sus scrofa*	NA	Japan	201	06-NOV-2015
LC067574	*Entamoeba polecki*	ST3	*Sus scrofa*	faecal sample	Japan	809	29-JUL-2015
KF646242	*Entamoeba polecki*	ST4	NA	NA	NA	474	30-MAR-2015
KF646243	*Entamoeba polecki*	ST4	NA	NA	NA	473	30-MAR-2015
KF646244	*Entamoeba polecki*	ST4?	NA	NA	NA	472	30-MAR-2015
EF110870	*Entamoeba polecki*	ST4	NA	NA	NA	478	09-DEC-2006
EF110871	*Entamoeba polecki*	ST4	NA	NA	NA	479	09-DEC-2006
EF110872	*Entamoeba polecki*	ST1	NA	NA	NA	477	09-DEC-2006
EF110873	*Entamoeba polecki*	ST3	NA	NA	NA	478	09-DEC-2006
EF110874	*Entamoeba polecki*	ST4	NA	NA	NA	479	09-DEC-2006

EF110875	*Entamoeba polecki*	ST3	NA	NA	NA	478	09-DEC-2006
EF110876	*Entamoeba polecki*	ST3	NA	NA	NA	478	09-DEC-2006
EF110877	*Entamoeba polecki*	ST4	NA	NA	NA	478	09-DEC-2006
EF110878	*Entamoeba polecki*	ST2	NA	NA	NA	478	09-DEC-2006
EF110879	*Entamoeba polecki*	ST2	NA	NA	NA	478	09-DEC-2006
EF110880	*Entamoeba polecki*	ST4	NA	NA	NA	478	09-DEC-2006
EF110881	*Entamoeba polecki*	ST4	NA	NA	NA	478	09-DEC-2006
DQ286374	*Entamoeba polecki*	ST3	pig	NA	NA	1036	03-DEC-2005
AF149913	*Entamoeba polecki*	ST1	NA	NA	NA	1858	22-DEC-1999
ON254800	*Entamoeba chattoni*	ST2	NHPs	NA	China	180	29-JUN-2022
MH133209	*Entamoeba chattoni*	ST2	rhesus macaques	faecal samples	China	579	01-APR-2019
MH626613	*Entamoeba chattoni*	ST2	NA	NA	NA	540	21-JUL-2018
MH620471	*Entamoeba chattoni*	ST2	NA	NA	NA	554	20-JUL-2018
KJ149294	*Entamoeba chattoni*	ST2	*Leontopithecus chrysomelas*	host faeces	United Kingdom	601	16-JUN-2014
KJ149295	*Entamoeba chattoni*	ST2	*Trachypithecus auratus auratus*	host faeces	United Kingdom	604	16-JUN-2014
AB749459	*Entamoeba chattoni*	ST2	*Macaca mulata*	NA	China	166	22-SEP-2012
AF149912	*Entamoeba chattoni*	ST2	*Papio cynomolgi*	NA	USA	1863	22-DEC-1999
MN192186	*Entamoeba struthionis*	ST3	NA	NA	Greece	627	11-DEC-2019
AB851494	*Entamoeba struthionis*	ST3	*Homo sapiens*	HIV patient	Cameroon	299	23-FEB-2016

AB851496	*Entamoeba struthionis*	ST3	*Homo sapiens*	HIV patient	Cameroon	294	23-FEB-2016
KF646239	*Entamoeba struthionis*	ST3	NA	NA	NA	386	30-MAR-2015
KF646240	*Entamoeba struthionis*	ST3	NA	NA	NA	451	30-MAR-2015
AJ566411	*Entamoeba struthionis*	ST3	*Struthio camelus*	intestinal content	Spain	1863	16-JUN-2003

Table 16. The host species of *Entamoeba bovis* identified so far based on DNA sequence data from the NCBI nucleotide database (as of January2023).

Host species (or information as provided in GenBank)	Number of individuals identified as hosts
Yak	284
Fallow deer	38
Bos taurus	18
Cattle	14
Deer	6
Sambar deer	6
Caprine	3
Aries ovis	2
Bovine	2
Ovine	2
Rangifer tarandus	1
Total	376

Table 17. *Endolimax*-specific SSU rDNA sequences currently available in the NCBI nucleotide database (January 2023) listed according to date of deposition.

NCBI Acc.no.	Organism	Host	Isolation source	Country	Sequence length	Date of deposition
OK483220.1	*Endolimax nana*	NA	Sewage	Sweden	1605	11-OCT-2022
OK483221.1	*Endolimax nana*	NA	Sewage	Sweden	1571	11-OCT-2022
OK483222.1	*Endolimax nana*	NA	Sewage	Sweden	1593	11-OCT-2022
OK483223.1	*Endolimax nana*	*Homo sapiens*	Stool	Denmark	1739	11-OCT-2022
MN508053.1	*Endolimax* sp.	NA	Waste water	Sweden	678	11-OCT-2022
MN508054.1	*Endolimax* sp.	NA	Waste water	Sweden	671	15-JAN-2020
MN508055.1	*Endolimax* sp.	NA	Waste water	Sweden	671	15-JAN-2020
MN508056.1	*Endolimax* sp.	NA	Waste water	Sweden	679	15-JAN-2020
LC230011.1	*Endolimax* sp. TDP-2	*Sus scrofa domesticus*	faeces of pig	Japan	2582	18-OCT-2019
LC230012.1	*Endolimax* sp. TDP-2	*Sus scrofa domesticus*	faeces of pig	Japan	2580	18-OCT-2019
LC230013.1	*Endolimax* sp. TDP-2	*Sus scrofa domesticus*	faeces of pig	Japan	2584	18-OCT-2019
LC230014.1	*Endolimax* sp. TDP-2	*Sus scrofa domesticus*	faeces of pig	Japan	2580	18-OCT-2019
LC230015.1	*Endolimax* sp. TDP-2	*Sus scrofa domesticus*	faeces of pig	Japan	2580	18-OCT-2019
MN556101.1	*Endolimax* sp.	*Homo sapiens*	stool	Denmark	1736	16-OCT-2019
MH979372.1	*Endolimax nana*	NA	Waste water	Australia	137	01-OCT-2018
MH979379.1	*Endolimax nana*	NA	Waste water	Australia	120	01-OCT-2018
MH979380.1	*Endolimax nana*	NA	Waste water	Australia	120	01-OCT-2018
MH979381.1	*Endolimax nana*	NA	Waste water	Australia	120	01-OCT-2018
MH979382.1	*Endolimax nana*	NA	Waste water	Australia	137	01-OCT-2018
MH979386.1	*Endolimax nana*	NA	Waste water	Australia	137	01-OCT-2018
MH979392.1	*Endolimax nana*	NA	Waste water	Australia	137	01-OCT-2018
MH979396.1	*Endolimax nana*	NA	Waste water	Australia	137	01-OCT-2018
MG925069.1	*Endolimax nana*	*Homo sapiens*	stool	Tunisia	302	14-FEB-2018
MH979397.1	*Endolimax nana*	NA	Waste water	Australia	137	01-OCT-2018

MG925070.1	*Endolimax nana*	*Homo sapiens*	stool	Tunisia	673	14-FEB-2018
MG925071.1	*Endolimax nana*	*Homo sapiens*	stool	Tunisia	712	14-FEB-2018
MG925072.1	*Endolimax nana*	*Homo sapiens*	stool	Tunisia	371	14-FEB-2018
MG925073.1	*Endolimax nana*	*Homo sapiens*	stool	Tunisia	300	14-FEB-2018
MG925074.1	*Endolimax nana*	*Homo sapiens*	stool	Tunisia	296	14-FEB-2018
MG925075.1	*Endolimax nana*	*Homo sapiens*	stool	Tunisia	672	14-FEB-2018
JX101953.1	*Endolimax piscium*	*Solea senegalensis*	fish muscle	Spain	2971	01-MAR-2013
JX101954.1	*Endolimax piscium*	*Solea senegalensis*	fish muscle	Spain	2971	01-MAR-2013
JX101955.1	*Endolimax piscium*	*Solea senegalensis*	fish muscle	Spain	2971	01-MAR-2013
AF149916.1	*Endolimax nana*	primate	faeces?	?	2589	22-DEC-1999

Table 18. *Iodamoeba*-specific SSU rDNA sequences currently available in the NCBI nucleotide database (January 2023) listed according to date of deposition.

NCBI Acc.no.	Organism	Host	Isolation source	Country	Sequence length	Date of deposition
MK801423	*Iodamoeba* sp. RL2	*Sus scrofa domesticus*	faeces	Germany	2181	13-SEP-2019
MK801428	*Iodamoeba* sp. RL2	*Sus scrofa domesticus*	faeces	Germany	2171	13-SEP-2019
MK801432	*Iodamoeba* sp. RL2	*Sus scrofa domesticus*	faeces	Germany	1945	13-SEP-2019
MK801434	*Iodamoeba* sp. RL2	*Sus scrofa domesticus*	faeces	Germany	2168	13-SEP-2019
MK801437	*Iodamoeba* sp. RL2	*Sus scrofa domesticus*	faeces	Germany	2170	13-SEP-2019
MK801438	*Iodamoeba* sp. RL2	*Sus scrofa domesticus*	faeces	Germany	2167	13-SEP-2019
MK801445	*Iodamoeba* sp. RL2	*Sus scrofa domesticus*	faeces	Germany	2176	13-SEP-2019
MK801447	*Iodamoeba* sp. RL2	*Sus scrofa domesticus*	faeces	Germany	2173	13-SEP-2019
MK801448	*Iodamoeba* sp. RL2	*Sus scrofa domesticus*	faeces	Germany	2170	13-SEP-2019
MK801454	*Iodamoeba* sp. RL2	*Sus scrofa domesticus*	faeces	Austria	2171	13-SEP-2019
MK801459	*Iodamoeba* sp. RL2	*Sus scrofa domesticus*	faeces	Romania	2171	13-SEP-2019
MK801463	*Iodamoeba* sp. RL2	*Sus scrofa domesticus*	faeces	Romania	2180	13-SEP-2019
MK801464	*Iodamoeba* sp. RL2	*Sus scrofa domesticus*	faeces	Germany	2169	13-SEP-2019
MH623073	*Iodamoeba* sp. ZOTU 3246	NA	waste water	Australia	115	24-MAR-2019
MH623069	*Iodamoeba* sp. ZOTU 2906	NA	waste water	Australia	115	24-MAR-2019
MG925076	*Iodamoeba* sp. RL1	*Homo sapiens*	stool	Tunisia	463	14-FEB-2018
JX158584	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158585	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158595	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158597	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158600	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158603	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158606	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013

JX158607	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158608	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158609	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158610	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158613	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158617	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158618	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158621	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158622	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158625	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158626	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158627	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158628	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158629	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158630	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158631	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158632	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158633	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158639	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158640	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158641	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158642	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158643	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158644	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158645	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158646	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158647	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013

JX158650	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158652	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158651	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158653	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158656	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158658	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158659	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158661	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158663	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
JX158679	uncultured *Iodamoeba*	gorilla	stool	Cameroon	697	04-FEB-2013
KC922216	uncultured *Iodamoeba*	cattle	stool	China	718	27-AUG-2014
KC922234	uncultured *Iodamoeba*	sheep	stool	China	718	27-AUG-2014
KC922238	uncultured *Iodamoeba*	sheep	stool	China	799	27-AUG-2014
KC922246	uncultured *Iodamoeba*	sheep	stool	China	718	27-AUG-2014
KC922268	uncultured *Iodamoeba*	sheep	stool	China	718	27-AUG-2014
KC922276	uncultured *Iodamoeba*	sheep	stool	China	798	27-AUG-2014
KC922300	uncultured *Iodamoeba*	sheep	stool	China	718	27-AUG-2014
JN635740	*Iodamoeba* sp. RL2	*Homo sapiens*	purified cysts from faeces	Cuba	2187	03-OCT-2011
JN635741	*Iodamoeba* sp. RL1	*Homo sapiens*	faeces	NA	2376	03-OCT-2011
JN635742	*Iodamoeba* sp. RL1	*Homo sapiens*	faeces	NA	2193	03-OCT-2011
JN635743	*Iodamoeba* sp. RL1	*Homo sapiens*	faeces	NA	509	03-OCT-2011
JN635744	*Iodamoeba* sp. RL2	*Sus scrofa*	primary culture	United Kingdom	1190	03-OCT-2011
JN635745	*Iodamoeba* sp. RL1	*Homo sapiens*	purified cysts from faeces	Thailand	1752	03-OCT-2011
JN635746	*Iodamoeba* sp. RL1	*Homo sapiens*	purified cysts from faeces	Thailand	1961	03-OCT-2011
JN635747	*Iodamoeba* sp. RL2	*Homo sapiens*	purified cysts from faeces	Cuba	257	03-OCT-2011
JN635748	*Iodamoeba* sp. RL2	*Homo sapiens*	purified cysts from faeces	Cuba	256	03-OCT-2011

JN635749	*Iodamoeba* sp. RL2	*Homo sapiens*	purified cysts from faeces	Cuba	252	03-OCT-2011
JN635750	*Iodamoeba* sp. RL2	*Homo sapiens*	purified cysts from faeces	Cuba	255	03-OCT-2011
JN635751	*Iodamoeba* sp. RL2	*Homo sapiens*	purified cysts from faeces	Cuba	256	03-OCT-2011