Genetic diversity of Entamoeba Species among Children Under 5 years in the Vhembe District, Limpopo Province, South Africa

Marc Maponya
University of Venda
Pascal Bessong
University of Venda
R Ngobeni
Cape Peninsula University of Technology
Ali Elbakri (aelbakri@sharjah.ac.ae)
University of Sharjah College of Health Sciences
Amidou Samie
University of Venda

Research

Keywords: Entamoeba species, Genetic diversity, Phylogeny, South Africa

Posted Date: December 15th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-125092/v1

Abstract

Background: Our understanding of the disease caused by the various *Entamoeba* species and its epidemiology is changing with time, and very little is known about their genetic diversity. Therefore, this study sought to investigate the prevalence and genetic diversity of *Entamoeba* species among children under 5 years in the Vhembe District, South Africa.

Methods: A total of 534 stool samples from 313 children (both males and females) aged under 5 were collected from 12 villages in Vhembe District, Limpopo. The prevalence of *Entamoeba* infections was examined by microscope and PCR, followed by Sanger sequencing for specific regions of 18S rRNA gene to identify and differentiate the circulating species of *Entamoeba* in the study population.

Results: Of the 313 children recruited in the study, 163 were females and 150 were males, the ages ranged from 1 to 3 years. Of the 534 samples, 130/534 (24.3%) were microscopically positive for *Entamoeba* cysts. However, *Entamoeba* genus-specific DNA amplification using PCR identified 43/534 (8%) of *Entamoeba* species. Twenty positive amplicons were sequenced by Sanger sequencing technologies. Out of the twenty samples, twelve (60%) were confirmed to be *Entamoeba* species. The *Entamoeba* species identified in the study as evidenced by BLAST calls in the NCBI database and phylogenetic tree after narrowing the search using the option *Entamoeba* taxid (5758) from the NCBI database included 4 *E. polecki* (33.3%), 6 *E. coli* (50%), one *E. muris* (8.3%) and one *E. hartmanni* (8.3%). The phylogenetic tree showed the close relationship between isolated species and the ones in the GenBank.

Conclusion:

The current study has shown for the first time the presence of *E. polecki* in humans and the existence of possibly two types of *E. coli* infecting humans. Our findings further emphasize the need for the re-evaluation of the pathogenicity of species such as *E. polecki* which are quite common in the study population and might be responsible for some of the health complications.

Background

Entamoeba genus contains a group of unicellular, anaerobic and parasitic microorganisms that infect the gastrointestinal tract of both humans and animals [1]. Despite the focus is largely due to *E. histolytica*, a major cause of morbidity and mortality in humans and animals [2], other species including *E. dispar, E. moshkovskii, E. coli, E. polecki, E. muris and E. hurtmani* are also found in this genus [3].

Entamoeba polecki, a uni-nucleated cyst-producing *Entamoeba* species, infections in humans are infrequent and are mainly linked with animal contact such as pigs [4–7]. Other reported cases of *Entamoeba* infections in human includes *Entamoeba chattoni*, a uninucleated cyst-producing *Entamoeba* species contracted during contact with monkeys [8]. However, whether they occur in humans or are even genetically distinct remains to be established [6].

The genetic diversity of *Entamoeba* species is one of the major studies that one can follow to understand the species in detail [9]. The identification of pathogenic species of *Entamoeba* may provide insight knowledge about the treatment, control, diagnosis and the epidemiology of the species [10]. Due to the humans and animals' medical purposes, the diversity of these parasites has been investigated [9]. Even though many *Entamoeba* species have been isolated and identified using molecular methods, still the genetic diversity of the *Entamoeba* species is poorly understood [9].

Studies of the diversity of *Entamoeba* species have been reported worldwide. Feng *et al*., (2018) reported a clear diverse of *Entamoeba* species in China in which the group highlighted the detection of 50% of *Entamoeba polecki* in pigs and humans' stool samples [11]. Data on the genetic diversity of *Entamoeba* species is scarce and only few studies have examined the genetic assortment of *Entamoeba* African strains. The present study sought to investigate the genetic diversity of *Entamoeba* species among children under 5 in the Dzimauli population, Vhembe District, Limpopo province in South Africa.

Methods

Study area, population and sample collection

The current study was carried out in 19 rural villages in Vhembe, Limpopo province, South Africa. In brief, the villages are inhabited by members of the indigenous ethic groups of Vha-Venda people. Almost all village members are still practicing their traditional culture. Majority of villagers live in deprived circumstances, where there is an overcrowding, poor sanitation, low levels of education and inadequate water supply. Rivers located adjacent to the villages remains the main source of water for domestic needs, such as drinking, cooking, bathing and washing clothes.

The environmental conditions of the villages are generally poor with limited provision of latrine facilities. Therefore, villagers especially children use the bucket toilet system and practice open defecation in close proximity to the rivers. Pigs, cattle, dogs, and chicken are the most common domestic animals found around the villages and some of the animals roam around freely and defecate in the surrounding property without supervision.

Microscopic examination

All stool samples were microscopically examined. A small amount of faecal material was mixed with a drop of iodine on a clean and dry microscope slide. The mixture was covered with a cover slip and observed under low (10×) and high (40×) magnification for the presence of *Entamoeba* cysts and other intestinal parasites. The presence of *Entamoeba* cysts in stool samples were identified based on their morphological characteristics and the number of nuclei observed.

Genus-specific PCR assay, sequencing and phylogenetic assay

Genomic DNA was extracted from all stool samples using a QIAamp DNA Stool Mini Kit (Qiagen) following manufacturer's instructions (Qiagen, Inc., Hilden, Germany). A PCR was performed using genus-specific PCR primers based on small-subunit rRNA gene sequences. Primer sequences (Entam1: 5´GTT GAT CCT GCC AGT ATT ATA TG 3´ and Entam2: 5´CAC TAT TGG AGC TGG AAT TAC 3´) which produce an approximate 550bp fragment length previously described by Verweij et al. (2001) were used [6]. Briefly, genus-specific PCR amplifications were performed in a final volume of 25µl using the thermal cycler (P100TM Thermal Cycler, BIO-RAD). For cycling conditions all reactions involved an initial denaturation step at 94 ˚C for 5 minutes followed by 35 cycles of 94 ˚C for 1 minute, 55 ˚C for 1 minute and 72 ˚C for 1 minute and a final extension at 72˚C for about 7 minutes. The PCR products were separated by electrophoresis in 2% agarose gel and visualized by UV-trans-illuminator.

For further analysis all positive amplicons were sequenced at a sequencing company (Inqaba Biotech, Pretoria, South Africa). Both strands were sequenced with the primers used for PCR. Sequenced data were aligned, analyzed and edited using BioEdit editor and the evolutionary relationship between the species was inferred using MEGA 10 software [12]. The SSU rRNA gene sequences obtained in this study were deposited in GenBank under accession numbers MW133761 - MW133772

Results

The mean age of the participants in the current study was 1.39 years. Five hundred and thirty four stool specimens were included to detect *Entamoeba* genus from symptomatic and asymptomatic children. Vegetative and/or cyst forms were found in 24.3% (130/534) by direct wet mount microscopy as *Entamoeba* cyst, either singly or in combination with other intestinal parasites such as *Endolimax nana* and *Trichuris trichura*. As for the genus specific PCR, 8% (43/534) were identified as positive for the *Entamoeba* genus (Fig. 1). Twenty were randomly selected for Sanger sequencing of which, twelve were confirmed to be *Entamoeba* species as evidenced by BLAST calls in the NCBI database and phylogenetic tree (Fig. 2) after narrowing the search using the option *Entamoeba* taxid (5758). The Blast calls result included: 4 *E. polecki* (33.3%), 6 *E. coli* (50%), one *E. muris* (8.3%) and one *E. hartmanni* (8.3%).

Discussion

To our knowledge, no study has investigated the prevalence and genetic diversity of *Entamoeba* species among children under 5 years in the Vhembe District, Limpopo, South Africa. In the present study, we have shown for the first time that the infection rate with *Entamoeba* species is 8% using molecular tools. Compared to PCR and sequencing, vegetative and/or cyst forms were found in 24.3% (130/534) by direct wet mount microscopy as *Entamoeba* cysts highlighting the trouble many laboratory technicians face in identifying and differentiating morphologically similar cysts and/or trophozoites of *Entamoeba* genus such as *E. histolytica*, and other uni-nucleated cysts including immature cysts of *E. histolytica* [14–16].

To investigate the genetic diversity of *Entamoeba* species in the study population, twenty samples were Sanger sequenced resulting in four *Entamoeba* species, *E. coli* (60%), *E. polecki* (33.3%), and *E. muris* and *E. hartmanni* (8.3%). Although the identified species might be less pathogenic in the case of a single infection, coinfections with other pathogens including bacterial, fungal and viral infection may augment the severity of the disease [17].

A large genetic distance exists between the un-, tetra-, and octanucleated cyst forming *Entamoeba* species as described by Silberman et al (1999) [18]. As presented in the phylogenetic tree, all *Entamoeba polecki* isolates clustered with *E. polecki* (AB845671) and *E. coli* (AB845674) reference sequence. On the other hand, they are widely separated from *E. coli* (AB444953), *E. muris* (FN396613) and the tetra-nucleated cyst forming *E. histolytica* and *E. moshkovskii*. Interestingly, possibly two variants of *E. polecki* are clearly distinguishable in the phylogenetic tree (Fig. 2). Isolate 46 is further away in the tree from the other three *E. polecki* samples suggesting that possibly two variants of *E. polecki* are identified in the present study. It has previously been proposed that variants of *E. polecki* exist since there is no host specificity and no known difference except for small amounts of sequence divergence [6, 19].

Sequencing revealed more *E. coli* positive samples clustered in two distant parts of the phylogenetic tree suggesting that they may be different species/types/stains of *E. coli*. Figure 2 shows that isolates 22, 24 and 27 clustered with *E. coli* (AB845674) reference sequence whereas isolates 28, 5 and 16 all clustered together and were widely separated from the other *E. coli* samples and reference sequence. Stensvold et al. (2011a) reported that *E. coli* samples from humans group into two clusters, which have been named subtypes 1 and 2 (ST1 and ST2) with ST1 widespread among humans [19]. Whether this variation is a result of the possible source of infection, human or animal origin, remains to be established. Sample 3, *Entamoeba hartmanni*, s tetra-nucleate cyst producing *Entamoeba* visibly branches out separately in the phylogenetic tree away from the other tetra-nucleated *Entamoeba* species (Fig. 2).

Studies done by Stensvold *et al.*, (2011b) demonstrated human infections with *E. polecki*, in which a novel 18S rRNA gene sequence was identified in a species of Sulawesi macaque [20]. However, in many cases, the local prevalence of these species may vary significantly based on the different geographical regions. A study done in South Africa reported that *E. polecki* (90%) were more prevalent as compared to *E. coli* with (10%) [21]. Furthermore, another study was reported in India which reported about 49.5% of *E. polecki* and only 7.4% with *E. coli and E. moshkovskii* [21]. *Entamoeba polecki* is mostly isolated from domestic animals especially pigs [22]. Therefore, looking at the study population setting we can also suggest that the infection might be transmitted from pigs to water than humans. Only one sample (#23) returned 100% identity with *E. muris* (FN396613) and *E. coli* (AB444953) suggesting the sample could either be infected from an animal or a human source. Both *E. muris* and *E. coli* are producers of octa-nucleated cysts and both look identical morphologically under the microscope.

Conclusion And Recommendations

The current study has shown for the first time the presence of *E. polecki* in humans and the existence of possibly two types of *E. coli* strains infecting humans. What is clear from this study, is that humans can undoubtedly be infected with uninucleated cyst-producing and that more genetic variability exists within this group as well as *E. coli* than has previously been recognized in human infections. Our findings further emphasize the need for the re-evaluation of the pathogenicity of species such as *E. polecki* which are quite common in the study population and might be responsible for some of the health complications. Several studies concerning the virulence factors and pathogenicity of the identified species still need to be done.

Declarations

Authors' contributions

MMM performed the experiments, analyzed and interpreted the data and drafted the paper. EA analyzed the data and modified the manuscript. NR, BPO and SA conceived and designed the experiments; analyzed and interpreted the data and modified the manuscript. All authors read and approved the final version of the manuscript.

Funding

This research was supported by National research foundation (NRF) as well as the Bill and Melinda Gates Foundation through the MAL-Ed TAC project. The other part of the funding came from the FNIH through the MADI trial project.

Availability of data and materials

All data supporting the conclusions are included in the article. Nucleotide sequences from the present study are available in the GenBank database under the accession numbers (MW133761 - MW133772).

Ethical clearance

The study was approved by the University of Venda Ethics Committee. The study details including reason for the survey and procedure of stool sample collection were explained to all parents of the children as well as the village leaders. Oral informed consent was obtained from all participants.

Competing interest statement

The authors declare no conflict of interest.

Consent for publication

Not applicable.

Additional information

No additional information is available for this paper.

References

1. Ngobeni N, 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;216:1592-1600.
2. Elsheikha HM, Regan CS, Clark CG. Novel *Entamoeba* findings in nonhuman primates. Trends Parasitol. 2018;34(4):283-94.
3. Hirashima Y, Manchanayake T, Yano T, Kitahara S, Koreeda T, Kamimura S, et al. Development of molecular diagnostic protocols for detecting three types of *Entamoeba* from diarrheal and asymptomatic pigs and environmental moist soils. Parasitol Res. 2017;116(7):2001-007.
4. Burrows RB, Klink GE. *Endamoeba polecki* infections in man. Am J Hyg. 1955;62: 156-67.
5. Chacin-Bonilla L, Bonilla E, Parra AM, Estevez J, Morales LM, Suarez H. Prevalence of *Entamoeba histolytica* and other intestinal parasites in a community from Maracaibo, Venuzuela. Ann Trop Med Parasitol. 1992;86:373-80.
6. Verweij JJ, Polderman AM, Clark Genetic Variation among Human Isolates of Uninucleated Cyst-Producing Entamoeba Species. J Clin Microbiol. 2001;39(4):1644-6. doi:10.1128/JCM.39.4.1644-1646.2001.
7. Desowitz RS, Barnish G. *Entamoeba polecki*and other intestinal protozoa in Papua New Guinea Highland children. Ann Trop Med Parasitol. 1986;80:399-402.
8. Sargeaunt PG, Patrick S, O'Keeffe D. Human infections of *Entamoeba chattoni*masquerade as *Entamoeba histolytica*.Trans. R Soc Trop Med Hyg. 1992;86: 633-34.
9. Kawano T, Imada M, Chamavit P, Kobayashi S, Hashimoto T, Nozaki T. 2017. Genetic diversity of *Entamoeba*: Novel ribosomal lineages from cockroaches. PloS one 2017;12(9):e0185233.
10. Partida-Rodríguez O, Serrano-Vázquez A, Nieves-Ramírez ME, Moran P, Rojas L, Portillo T, et al. 2017. Human intestinal microbiota: interaction between parasites and the host immune response. Arch Med Res. 2017; 48(8):690-700.
11. Feng Y, Ryan UM, Xiao L. Genetic diversity and population structure of Cryptosporidium. Trends Parasitol. 2018;34(11):997-1011.
12. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Mol Biol Evol. 2018;35:1547-49.

13. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol. 1993;10:512-526.
14. Van Wyk JA, Mayhew E. 2013. Morphological identification of parasitic nematode infective larvae of small ruminants and cattle: A practical lab guide. Onderstepoort J Vet Res. 2013;80(1):23-30.
15. Kaya D, Demirezen Ş, Hasçelik G, Kıvanç DG, Beksaç MS. 2013. Comparison of PCR, culturing and Pap smear microscopy for accurate diagnosis of genital Actinomyces. J Med Microbiol. 2013;62(5):727-33.
16. Stensvold CR, Winiecka-Krusnell J, Lier T, Lebbad M. 2018. Evaluation of a PCR method for detection of *Entamoeba* polecki, with an overview of its molecular epidemiology. J Clin Microbiol. 2018;56(5):e0015418.
17. Fletcher SM, Stark D, Harkness J, Ellis J. Enteric protozoa in the developed world: a public health perspective. Clinical microbiology reviews, 2012;25(3):420-49.
18. Silberman JD, Clark CG, Diamond LS, Sogin ML. (1999). Phylogeny of the Genera Entamoeba and Endolimax as Deduced from Small-Subunit Ribosomal RNA Sequences. Mol Biol Evol. 1999;16:1740-51.
19. Stensvold CR, Lebbad M, Victory EL, Verweij JJ, Tannich E, Alfellani M, et al. Increased sampling reveals novel lineages of *Entamoeba*: consequences of genetic diversity and host specificity for taxonomy and molecular detection. 2011a; 162:525-41.
20. Stensvold CR, Lebbad M, Verweij JJ. The impact of genetic diversity in protozoa on molecular diagnostics. Trends in Parasitol. 2011b;27(2):53-58.
21. Sylvain PN, Kaur U, Goyal K, Sehgal R, Paul MF. Molecular differentiation of *Entamoeba* isolated from Cameroonian human immunodeficiency virus (HIV) infected and uninfected patient. J Parasitol Vector Biol. 2015;7:139-50.
22. Matsubayashi M, Sasagawa Y, Aita T, Tokoro M, Haritani, Shibahara T. First report of mixed *Entamoeba polecki*(ST 1) and *suis* infection in piglets shedding abnormal feces by histopathological and molecular surveys. Acta Parasitol. 2016; 61(4):665-70.

Figures

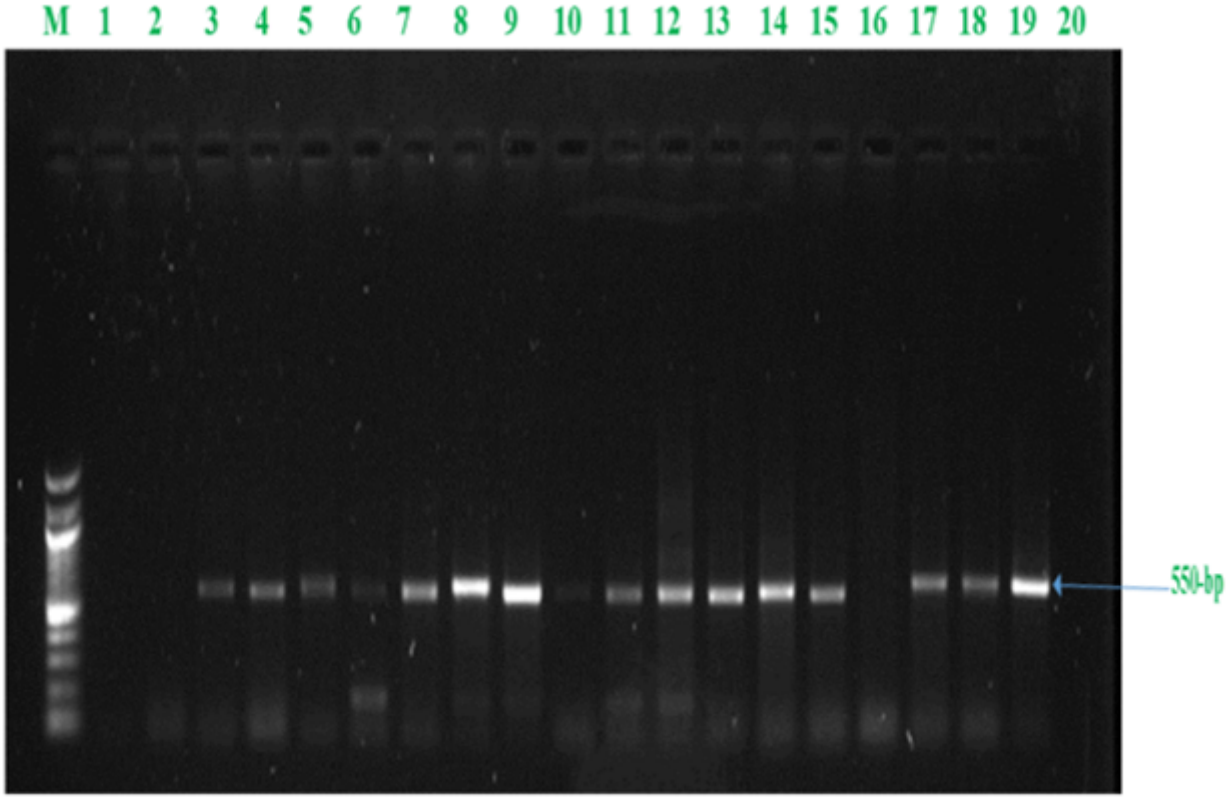

Figure 1

PCR amplification of Entamoeba DNA using genus-specific primers. Lane M= molecular marker (100-bp), Lane 1= negative control (PCR water), Lanes 2 to 20 = amplified products (550bp) indicating positive specimens.

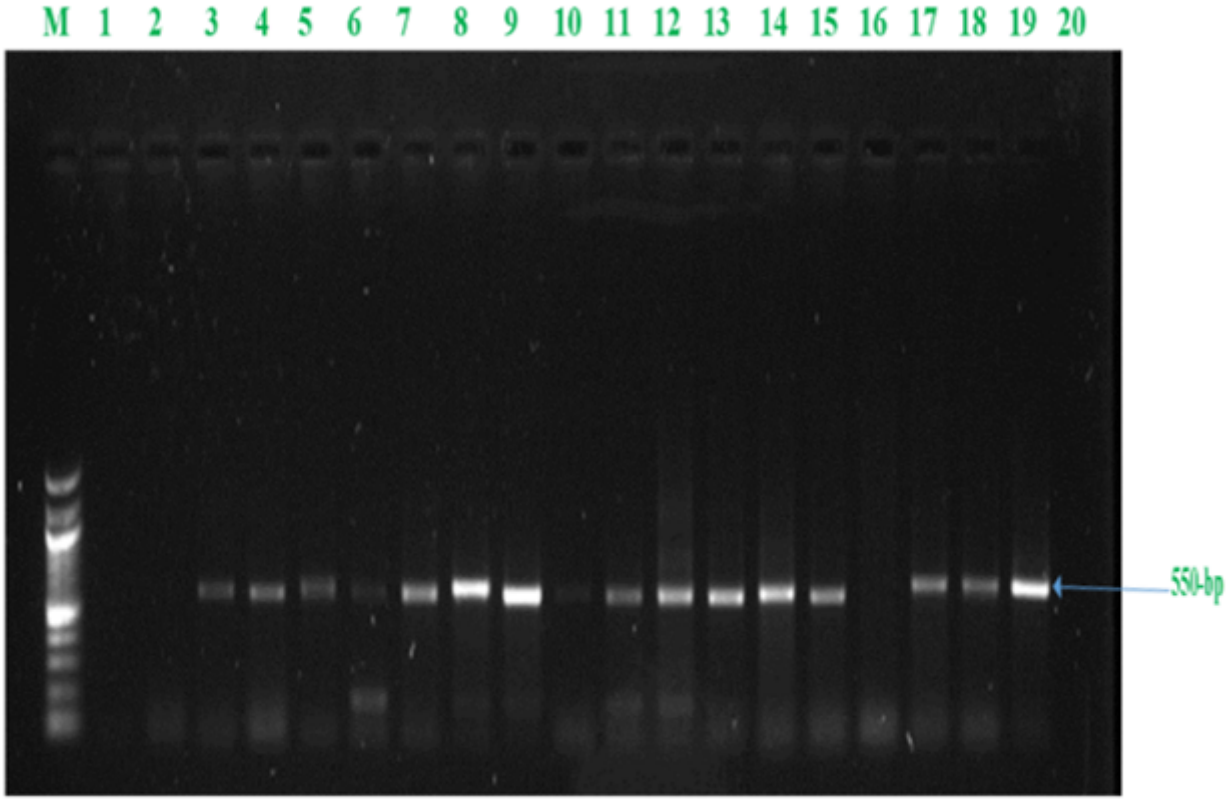

Figure 1

PCR amplification of Entamoeba DNA using genus-specific primers. Lane M= molecular marker (100-bp), Lane 1= negative control (PCR water), Lanes 2 to 20 = amplified products (550bp) indicating positive specimens.

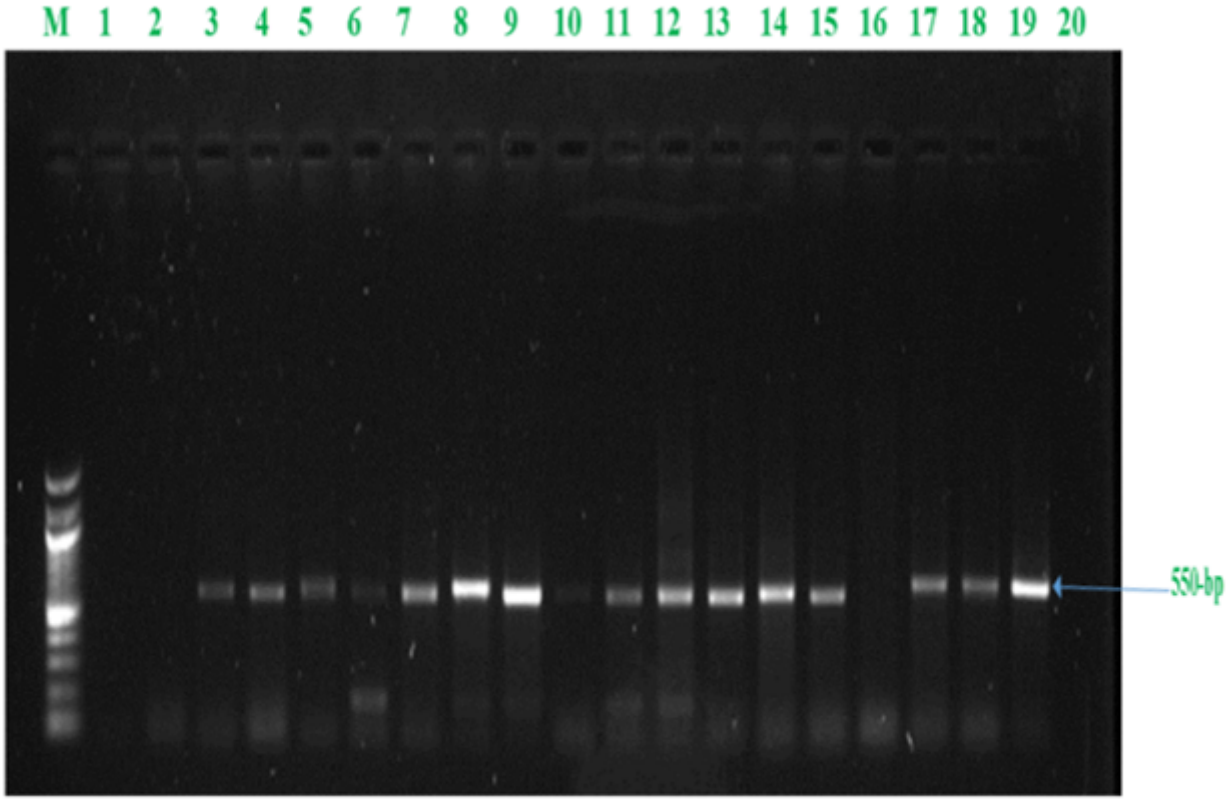

Figure 1

PCR amplification of Entamoeba DNA using genus-specific primers. Lane M= molecular marker (100-bp), Lane 1= negative control (PCR water), Lanes 2 to 20 = amplified products (550bp) indicating positive specimens.

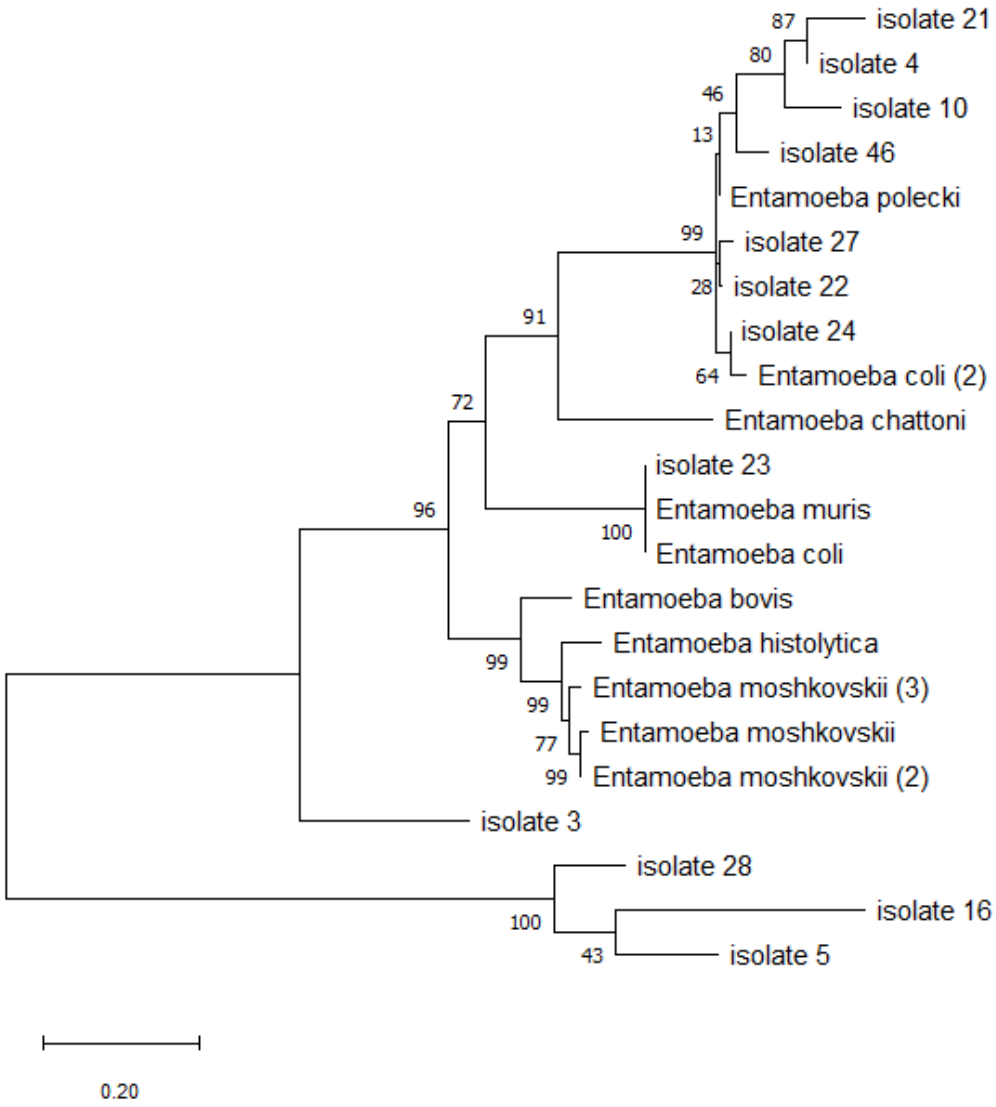

Figure 2

The evolutionary history was inferred by using the Maximum Likelihood method and Tamura-Nei model [13]. The tree with the highest log likelihood (-6058.62) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the

topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 22 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. There were a total of 695 positions in the final dataset. Evolutionary analyses were conducted in MEGA X (Kumar et al., 2018)

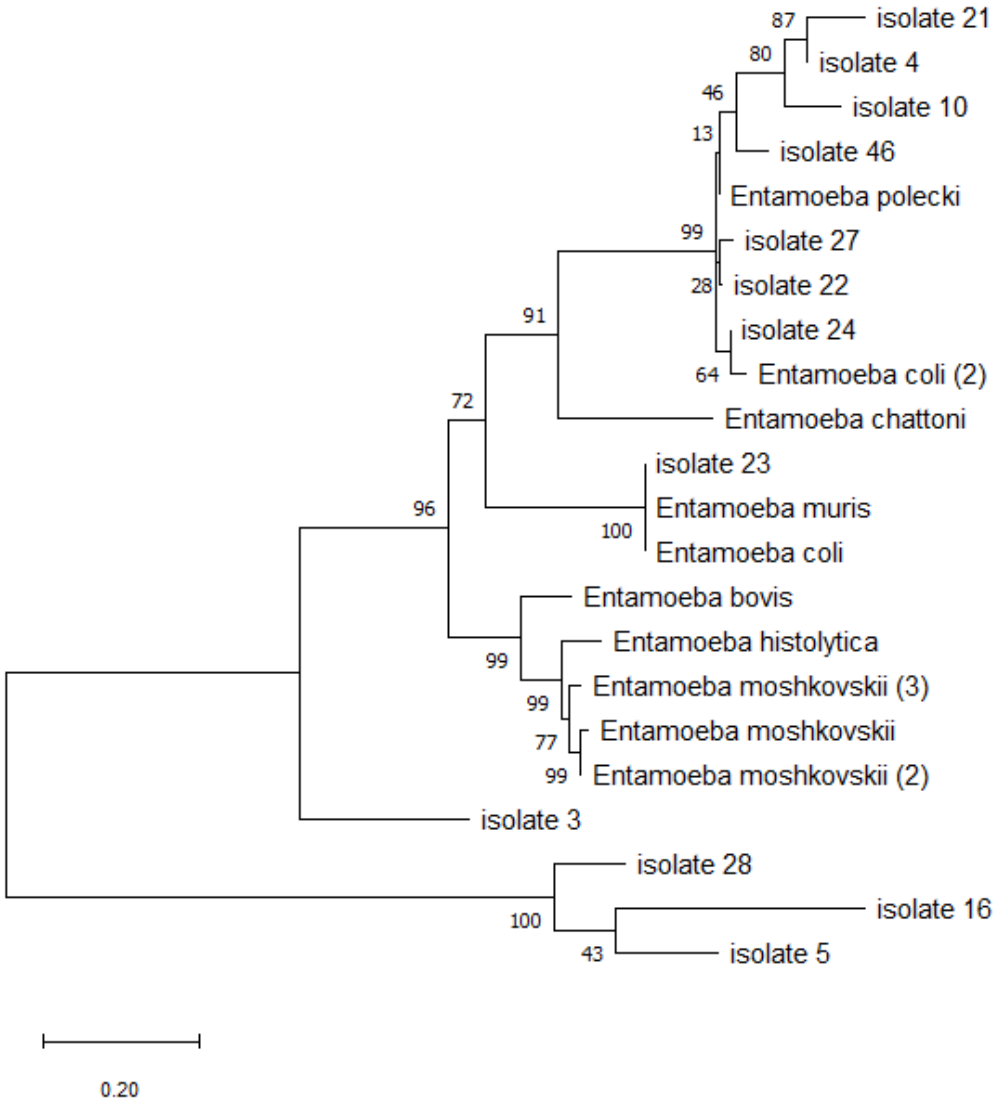

Figure 2

The evolutionary history was inferred by using the Maximum Likelihood method and Tamura-Nei model [13]. The tree with the highest log likelihood (-6058.62) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 22 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. There were a total of 695 positions in the final dataset. Evolutionary analyses were conducted in MEGA X (Kumar et al., 2018)

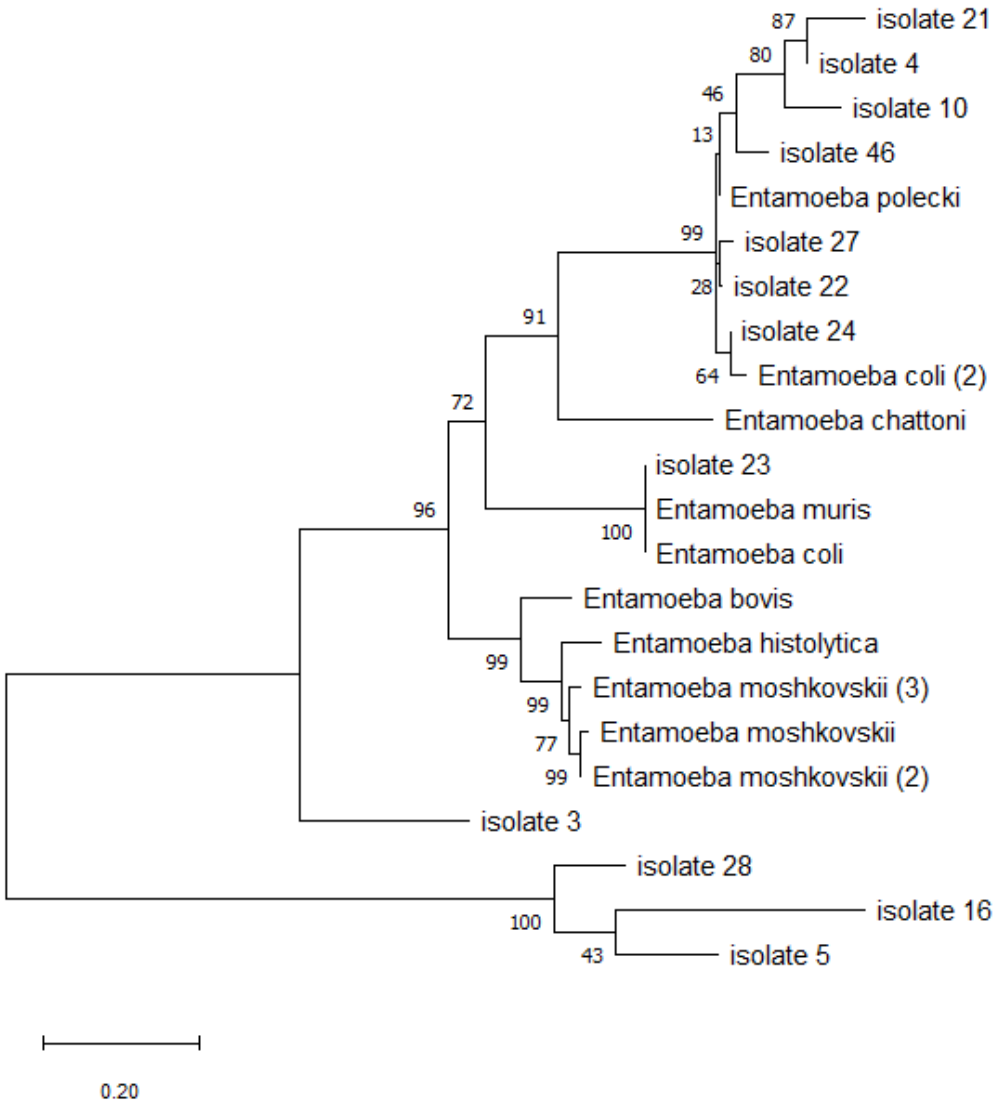

Figure 2

The evolutionary history was inferred by using the Maximum Likelihood method and Tamura-Nei model [13]. The tree with the highest log likelihood (-6058.62) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the

topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 22 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. There were a total of 695 positions in the final dataset. Evolutionary analyses were conducted in MEGA X (Kumar et al., 2018)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Graphicalabstract.docx
- Graphicalabstract.docx
- Graphicalabstract.docx