Review

Novel *Entamoeba* findings in non-human primates

Hany M. Elsheikha,[1] Carl S. Regan,[1,2] and C. Graham Clark [3, *]

[1] Faculty of Medicine and Health Sciences, School of Veterinary Medicine and Science, University of Nottingham, Sutton Bonington Campus, Loughborough, LE12 5RD, UK

[2] Current address: Vets4pets Dover Whitfield, White Cliffs Retail Park, Whitfield, Dover, CT16 3PS, UK

[3] Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, UK

*Correspondence:
graham.clark@lshtm.ac.uk
(C. Graham Clark)

Keywords: *Entamoeba*, Evolution, Non-human primates, Phylogeny, Species delineation

Abstract

In addition to well-known human-infecting species, *Entamoeba* species not found in humans have been identified recently in non-human primates (NHPs). Importantly, it has become clear that the organism identified as *Entamoeba histolytica* in NHPs is usually a distinct species, *Entamoeba nuttalli*. Many DNA-based stool surveys use species-specific detection methods and so may miss the full range of *Entamoeba* species present. In addition, authors may be using the same species name to describe distinct organisms. These various shortcomings may not be obvious to readers. In this review, we clarify the relationships between *Entamoeba* species' names based on morphological and molecular data, and highlight gaps in recently published data on *Entamoeba* species in wild NHPs resulting from the use of variable methodology.

Humans and NHPs are both primates, but how similar are their *Entamoeba* species?

Humans are primates, and therefore it would be logical to assume that the parasite fauna of humans and **non-human primates** (NHPs; see Glossary) is likely to be similar. However, this simplistic view ignores the huge range of life-styles, diets and ecological specialisations exhibited by NHPs, and the millions of years of independent evolution that separate us from even our closest NHP relatives, the great apes. Nevertheless, humans and NHPs do appear to have many parasites in common, at least when identified via microscopy. Over recent decades, molecular tools have allowed us to re-examine these similarities and to challenge the assumption that apparent morphological identity equates to species identity. This review discusses how molecular tools provide a clearer picture of the relationships between intestinal amoebae of the genus *Entamoeba* in humans and NHPs and where gaps in our understanding remain.

What is causing invasive amoebiasis in humans and NHPs?

The focus on *Entamoeba* is largely due to *Entamoeba histolytica* being a significant cause of morbidity and mortality in humans. Published estimates suggest this organism is responsible for millions of cases of disease and over 50,000 deaths in humans annually [1]. Although these numbers are extrapolated from a limited number of studies, *E. histolytica* is certainly responsible for a significant amount of disease in some locations. Captive NHPs occasionally die from a disease that is, superficially, indistinguishable from that caused by *E. histolytica* in humans (e.g. [2]). Several other *Entamoeba* species that resemble *E. histolytica* morphologically have been described in both humans and NHPs, making microscopic diagnosis problematic. Morphologically distinct, non-pathogenic species of *Entamoeba* also appear to be shared by humans and NHPs, further complicating diagnosis (see below).

The morphology era

The existence of species of *Entamoeba* in humans and NHPs that appear identical by microscopy has been known for over a century. At that time, organisms in new hosts were often given new species names, whether morphologically distinguishable or not. A major work by Dobell [3] concluded that all named intestinal species of *Entamoeba* in humans could be assigned to either *E. histolytica* or *Entamoeba coli*, but he equivocated about *Entamoeba* from NHPs on the grounds of insufficient data; he later concluded that intestinal *Entamoeba* species in NHPs were also *E. histolytica* and *E. coli* [4]. His two-species nomenclature stayed essentially intact for 35 years.

In the mid-1950s, Burrows [5] resurrected the name *Entamoeba hartmanni* for an organism that parasitologists were referring to as 'small race *E. histolytica*'. Dobell [3] had viewed *E. hartmanni* as a synonym of *E. histolytica*; however, Burrows showed that the sizes of *E. histolytica* 'large race' and 'small race' cysts were not a continuum but had a clear bimodal distribution. This first 'break' with the Dobell nomenclature was quickly adopted, because parasitologists were already primed to accept it.

The molecular era

Entamoeba hartmanni was the last change to Dobell's nomenclature scheme based on morphology alone. Additional changes followed but not for many years, as the changes were primarily dependent on **small subunit ribosomal RNA gene** (SSU-rDNA) analyses. Emile Brumpt [6] proposed the existence of *Entamoeba dispar*, a non-pathogenic species morphologically identical to *E. histolytica*. This proposal was rejected by most parasitologists at the time (see discussion following [7]) and the name *E. dispar* virtually disappeared from the literature. Suspicion that Brumpt had been correct followed on from studies based on both lectin agglutination [8] and **isoenzyme** patterns [9], in which two groups within *E. histolytica* were identified, only one of which was found in patients with invasive disease. Subsequently, studies (cited in [10]) using monoclonal antibodies, DNA hybridization, SSU-rDNA restriction fragment length polymorphism, and eventually DNA sequencing all identified the same two groups of strains, and this led to the formal redescription of *E. dispar* as a species distinct from *E. histolytica* [10].

Other SSU-rDNA-based changes to the nomenclature of human *Entamoeba* species include the reassignment of '*E. histolytica*-like' amoebae to the species *Entamoeba moshkovskii* [11]

and the recognition that uninucleate cysts occasionally seen in humans were not always immature *E. histolytica* but were in fact *Entamoeba polecki* [12]. Most recently, *Entamoeba bangladeshi* was described as a new human species [13]; if it were not for SSU-rDNA sequences this organism would have been identified as *E. moshkovskii* despite it being genetically quite distinct.

The nomenclature for *Entamoeba* species in NHPs has followed suit, for the most part. *Entamoeba hartmanni* is commonly found in NHPs. *Entamoeba dispar* is also widespread in NHPs. *Entamoeba chattoni* had long been accepted as a NHP-specific species of *Entamoeba* with uninucleate cysts. It was designated a **subtype** of *E. polecki* a few years ago [14], but this change of nomenclature for *E. chattoni* has not been universally accepted; this will be discussed further below.

Thus, for the most part, the NHP *Entamoeba* nomenclature changes simply mirrored those in humans without any investigations to evaluate whether they were in fact the same organisms. This was understandable initially because there was no reason to suspect there were differences and the investigative tools were not readily available to many researchers. However, now that molecular techniques are routine in most research laboratories and some diagnostic laboratories, investigations into the diversity and identity of *Entamoeba* in NHPs have become more common and are revealing some surprising and important findings.

The evidence for *Entamoeba* genetic diversity in NHPs is based almost exclusively on SSU-rDNA analyses. Analyses of other markers are rarely possible because most studies use DNA extracted directly from stool samples, but when available they show the same species relationships. SSU-rDNA is a multicopy gene, which makes it relatively easy to amplify from

stool samples. In addition, and in contrast to some eukaryotes, the *Entamoeba* SSU-rDNA is relatively fast evolving (as evidenced by long branches in phylogenetic trees) meaning that sufficient resolution is obtained to allow differentiation of *Entamoeba* taxa using this gene alone.

Entamoeba nuttalli

Entamoeba histolytica causes disease of two main types: 1. amoebic dysentery/colitis, resulting from trophozoite invasion of the colonic mucosa and leading to ulceration, bleeding and the production of loose stool with blood and mucus; 2. amoebic liver abscess, resulting from haematogenous spread of trophozoites from the colon via the portal system to the liver, where tissue lysis leads to formation of a sterile pus-filled abscess [15]. Both types of disease have been reported in NHPs, and there have been a number of reports over the years of spontaneous invasive disease occurring in captive NHPs. Histologically, the diseases in humans and NHPs appear identical, as do the amoebae under the microscope [e.g. 2]. *Entamoeba histolytica* of human origin has been shown experimentally to be capable of infecting NHPs, where it can cause indistinguishable pathology [e.g. 16]. The organism responsible was therefore presumed to be *E. histolytica* in all cases of disease in NHPs.

In the last 10 years, however, molecular studies have been performed on amoebae from cases of invasive amoebiasis occurring spontaneously in NHPs. The amoebae in NHPs are consistently distinguishable from *E. histolytica* using a variety of DNA and protein markers: isoenzymes, SSU-rDNA and short tandem-repeat-containing loci [17-20]. Although closely related to *E. histolytica* – indeed it has been called "*E. histolytica*-like variant" [19] and "*E. histolytica* NHP variant" [21] by some – this is clearly a distinct organism and the name *E. nuttalli* has been revived for this amoeba [17]. *Entamoeba nuttalli* was originally described by

149 Castellani [22] in the liver abscess of a toque macaque (*Macaca sinica*) in Sri Lanka and is one
150 of the species considered synonymous with *E. histolytica* by Dobell [3, 23]. Although we
151 cannot prove after 110 years that the amoeba observed by Castellani is the same as the one now
152 being called *E. nuttalli*, this seems quite likely. A recent survey of wild toque macaques in Sri
153 Lanka detected asymptomatic carriage of *E. nuttalli* in 18.5% of the 227 animals studied [24].
154 *Entamoeba histolytica* was not detected in the population. *Entamoeba nuttalli* has been found
155 in a variety of other NHPs – guenon, baboon, colobus and chimpanzee – in addition to other
156 species of both captive and wild macaques [19, 25, 26].
157
158 The host and geographic ranges of *E. nuttalli* seem to be quite large, but so far it seems to be
159 found primarily in primates of the Old World. Invasive disease has been reported in captive
160 spider monkeys [25], but whether it infects wild New World NHPs is unknown. Only one
161 human infection with *E. nuttalli* has been reported to date, in a zookeeper [27]. This is despite
162 analyses of human samples that would have revealed its presence if it had been there.
163 Isoenzyme analysis, which was used widely for *Entamoeba* species differentiation in the 1980s
164 and early 1990s [e.g. 28], would have distinguished *E. nuttalli* from *E. histolytica* [17, 19], but
165 although many thousands of human samples were studied in order to differentiate *E. dispar*
166 and *E. histolytica*, no evidence of what is now being called *E. nuttalli* was reported. A second
167 human infection has apparently been identified in Iraq, based on a sequence in GenBank stated
168 to be of human origin [55; GenBank accession number: KP233837].
169
170 Note that most DNA-based diagnostic tools cannot distinguish *E. nuttalli* from *E. histolytica*,
171 unless combined with sequencing, and neither can some commercial antigen-based diagnostic
172 kits and monoclonal antibodies [17]. Therefore, although it seems unlikely that significant
173 numbers of humans will be found to be infected with *E. nuttalli*, such infections may occur

174 occasionally among those who have close contact with NHPs, and may go unrecognized
175 depending on the diagnostic method used. Primer pairs specific for *E. nuttalli* do now exist [17,
176 25] so that positive identification of this species without sequencing is possible.
177
178 NHPs can be infected experimentally with *E. histolytica* cysts of human origin [23, 29],
179 although no invasive disease has resulted from such experiments. Captive NHP infections
180 involving *E. histolytica* have been confirmed by DNA sequencing [30]. Therefore, it cannot be
181 ruled out that some natural *E. histolytica* infections will occur in wild NHPs – most likely
182 among those that come into contact regularly with humans or human waste – although there
183 are no sequence-confirmed infections to date [e.g. 56]. It is impossible retrospectively to know
184 which organism was responsible for the invasive amoebiasis cases in NHPs reported in the
185 literature. Indeed, it is not possible to be certain that the amoeba observed was responsible for
186 the disease in some cases – the presence of an *Entamoeba* and dysentery in the same host does
187 not necessarily imply cause and effect.
188
189 ***Entamoeba polecki***
190 *Entamoeba polecki* produces cysts with one nucleus, as does *E. chattoni*. Sequencing of their
191 SSU-rDNAs revealed them to be closely related organisms [31]. The former species is
192 traditionally associated with pigs and the latter with NHPs. Despite sporadic reports of *E.*
193 *polecki* infections in humans for many years [32], when uninucleated cysts were seen in
194 humans it was generally assumed that they represented immature cysts of *E. histolytica* rather
195 than of *E. polecki* or *E. chattoni*. Verweij et al. [12] studied human *Entamoeba* infections where
196 only uninucleated cysts were seen and found four distinct SSU-rDNA sequences. Two of these
197 sequences were essentially identical to those of *E. polecki* and *E. chattoni* isolated from a pig
198 and a monkey, respectively, while the other two sequences were related but distinct. This meant

199 that there were four closely-related organisms with two names between them and that *E.*
200 *polecki* and *E. chattoni* were not host-specific since all four organisms were found in humans.
201
202 Verweij et al. proposed [12] that the four should be viewed as variants of the same organism
203 and called '*E. polecki*-like', as the name *E. polecki* has precedence. Later, Stensvold et al. [14]
204 proposed that they should be considered subtypes and numbered ST1-ST4, with the former *E.*
205 *polecki* becoming *E. polecki* ST1 and the former *E. chattoni* becoming *E. polecki* ST2. The
206 rationale for this approach is that there is no host specificity and no known difference except
207 for small amounts of sequence divergence. This subtype nomenclature has not been fully
208 accepted. One of the two 'unnamed' subtypes was in the interim named *Entamoeba struthionis*
209 [33] as it was isolated from an ostrich, but this subtype (ST3) has subsequently been found in
210 pigs [34] as well as humans. The fourth subtype has never had a species name and for a long
211 time was only known from humans, where it is the most common subtype. Recently, however,
212 ST4 was found to be the only *E. polecki* subtype in wild Celebes crested macaques (*Macaca*
213 *nigra*) [35], proving that *E. polecki* ST2 (*E. chattoni*) is not the only subtype found in NHPs.
214 It is possible that *E. polecki* ST1 and ST3 will also eventually be identified in NHP hosts. In
215 the absence of host-specificity, use of the '*E. polecki* subtype' nomenclature seems appropriate.
216
217 ***Entamoeba dispar*, *Entamoeba hartmanni* and *Entamoeba coli***
218 For the most part, these three species meet the original expectation that human and NHP
219 *Entamoeba* species are the same. *Entamoeba dispar* is quite a homogeneous species and there
220 is no indication to date that *E. dispar* from humans is in any way distinct from that in NHPs.
221 Although *E. hartmanni* shows a greater degree of SSU-rDNA variation than *E. dispar*, there is
222 no obvious clustering of sequences that reflects human or NHP origin [14, 36], suggesting it is
223 a discrete species with moderate intraspecific variation.

The situation in *E. coli* is more complex and less clear-cut. *Entamoeba coli* samples from humans group into two clusters, which have been named ST1 and ST2 [14]; ST1 appears to be slightly more common than ST2 in humans. When NHP *E. coli* samples are examined, the same two STs are identified, with ST2 being slightly more common, although this is based on relatively few samples. ST2 was recently identified in wild mountain gorillas (*Gorilla beringei*) [36]. The degree of divergence between the SSU-rDNAs of the two subtypes is substantial and distinct species names could be justified. However, other than this sequence divergence, there are no known differences between the two subtypes to date. *Entamoeba coli* cysts can vary quite dramatically in size [37, 38]. Whether this size variation is a morphological reflection of the underlying sequence divergence remains to be established.

Another *Entamoeba* that has been detected in NHPs is *Entamoeba* RL7 [14]. No species name has been assigned to this organism – it is simply known by its **ribosomal lineage** (RL) number [14]. *Entamoeba* RL7 was originally identified in a sample from a Phayre's leaf monkey (*Trachypithecus phayrei*) [14], but it has subsequently been detected in humans in West Africa [34]. Uniquely, this *Entamoeba* is most closely related to *Entamoeba muris* (Figure 1), which, like *E. coli*, produces cysts with eight nuclei. Based on morphology, this organism previously would have been reported as *E. coli*.

NHP-restricted *Entamoeba* Species

There are several NHP-restricted *Entamoeba* sequences worthy of discussion here. The first is *Entamoeba* RL3, which to date has only been detected in langurs of various species and produces cysts with a single nucleus. In the past it would likely have been reported as *E.*

chattoni based on microscopy. No infections with this organism have been reported in humans, or indeed in any other NHP. It is closely related to, but distinct from, *Entamoeba bovis* and related lineages that are confined to ungulates [14]. RL3 has only been found in a few samples but it is notable that two lineages of *Entamoeba* (RL3 and RL7) have to date been detected exclusively in langurs. Whether this is linked to their unusual foregut fermentative digestion is unclear.

Villanueva-García et al. [39] recently reported SSU-rDNA sequences of an apparently novel *Entamoeba* in two species of Howler monkey. Because these were only partial sequences they were given a **conditional lineage** identifier [34] rather than a RL number. *Entamoeba* CL8 is clearly distinct from previously sequenced *Entamoeba* SSU-rDNAs and, interestingly, the CL8 sequence branches within a cluster of sequences obtained from reptilian *Entamoeba* strains. Villanueva-García et al. found a second *Entamoeba* sequence in their samples that is virtually identical to *Entamoeba* RL6, which was originally described from the green iguana (*Iguana iguana*) [14, 40]. The complete SSU-rDNA sequence of both these organisms would be helpful in order to confirm their phylogenetic tree placement.

Finally, there has been one report of *Entamoeba suis* from a gorilla (*Gorilla gorilla*) [14], but whether this is a natural host for this *Entamoeba* species remains to be established. This species also produces cysts with a single nucleus.

Missing *Entamoeba* Species?

Perhaps surprisingly, there are to date no reports of *E. moshkovskii* from NHPs. This organism is actually a species complex with substantial intra-specific sequence variation [40] and is

being reported from humans with increasing frequency now that PCR-based detection is being employed [e.g. 41-43]. *Entamoeba moshkovskii* has also been detected in cattle, elephants [34], reptiles [44] and insects [Silberman JD, personal communication], so it is likely only a matter of time before it is also found in NHPs. Not all published molecular studies have tested for this species and in those that did it is not clear whether the primers used would detect all variants of this genetically diverse species complex. The most recently described *Entamoeba* of humans, *E. bangladeshi* [13], is also yet to be reported from NHPs.

Entamoeba gingivalis, which colonises the gingival pockets in the mouth of humans, is listed as having been found in NHPs [e.g. in 45]. No molecular data are available to know whether the organisms reported in NHPs differ from those in humans. This may be important, as there are at least two SSU-rDNA variants of *E. gingivalis* in humans [40] and additional diversity could exist in other hosts.

A summary of the relationships between species names and identifiers can be found in Table 1 and an outline phylogenetic tree depicting the relationships between *Entamoeba* SSU-rDNA sequences is depicted in Figure 1.

Captive vs. Wild NHPs

Data on the presence and prevalence of *Entamoeba* species in NHPs is patchy at best, and most reports are based on animals in zoological parks. This is a problem when it comes to interpreting the data. The first issue is how to interpret the presence of parasites in captive NHPs. Animals in captivity may be exposed to organisms they would never encounter in the wild. Therefore, the data only indicate that the NHP species is capable of becoming colonised

by the parasite identified, not that it is a natural host for this parasite. A second issue is the impact of captivity on prevalence. It is likely that animals come in contact with faeces and faecal contamination of food and water more frequently in captivity than they would in the wild; this is especially true of species that are primarily or exclusively arboreal. Only by studying wild NHPs can 'natural' infections be identified, although in the case of peri-urban and urban NHPs the possibility of infection through contact with human faeces cannot be excluded. It is, of course, also likely that wild NHPs will ingest faeces from other hosts, accidentally or on purpose. If the ingested faeces contains *Entamoeba* cysts it is possible that DNA of these organisms will be detected when the NHP faeces is screened by PCR. However, unless the NHP species ingests faeces frequently and in significant amounts it would be unlucky if the small amount of NHP faeces analysed contained detectable DNA of *Entamoeba* cysts that were just passing through.

Relatively few studies of *Entamoeba* in wild NHPs have employed molecular diagnostics to date, and microscopy does not differentiate most of the known *Entamoeba* species: only *E. histolytica*, *E. coli* and *E. chattoni* are regularly reported in publications reliant on microscopy. Each of these names actually represents a mixture of distinct organisms united only by the number of nuclei in their mature cyst. *Entamoeba hartmanni* is the only additional species that can be identified by morphology, but only if cyst diameters are measured; often this is either not the case or the information is not given. As a result, only studies employing sequence-based identification will be discussed below. We recognise that this excludes the vast majority of studies, but if the data are not interpretable we feel they are better omitted.

Molecular studies in wild NHPs published to date (Table 2) are few in number, mostly involve Old World NHPs, and vary in the methodology used. In some studies, species-specific PCR

has been used, but often not all known species were tested for despite primers being available, leaving gaps in the data (Table 2, notes). When species-specific PCR has been used, this often means subtypes were not identified and potentially interesting data on sequence variation and host range have been lost. Several studies did not test for *E. hartmanni*, leading to a false impression of the distribution of this *Entamoeba* species in NHPs. It is notable that *E. histolytica* was not detected in any of these studies.

The use of only species-specific primers can mean that novel *Entamoeba* species are missed. For example, if Villanueva-García et al. [39] had used species-specific primers for *Entamoeba*, the two novel *Entamoeba* species found in Howler monkeys (CL8 and RL6) would not have been identified – the samples would have appeared negative even though *Entamoeba* organisms were present. Sequencing of products amplified using genus-specific primers may seem the best way forward, but there is a catch. NHPs are often carriers of multiple *Entamoeba* species and mixed PCR products give unreadable sequences with the standard DNA sequencing. The approach of Jirků-Pomajbíková et al. [46] could be a good compromise – genus-specific amplification coupled with nested species-specific PCR. This allows identification of species in mixed infections yet does not miss mono-infections with novel *Entamoeba* species, as these would be positive with genus-specific but negative with all the species-specific primers used. Jirků-Pomajbíková et al. [46] did not initially test for *E. hartmanni* but through sequencing discovered that it was the *Entamoeba* present in the samples positive with the genus-specific primers but negative with the species-specific primer pairs used. However, this method will only identify the presence of novel *Entamoeba* species if they are present as a single infection unless it is combined with cloning of the PCR products.

It seems likely that identification of *Entamoeba* in NHPs in the future will be through microbiome data, whether from targeted amplification and sequencing of a portion of eukaryotic SSU-rDNA or by extraction of such sequences from metagenomic data. Both approaches are in use in humans and have identified *Entamoeba* when present, but to date have rarely been applied to NHP samples. In one example, Wegener Parfrey et al. [47] identified *E. hartmanni* (among many other eukaryotes) in captive NHPs through eukaryote-targeted SSU-rDNA amplification and 454 sequencing. Similarly, random sequencing of faecal DNA has the potential to identify not only all the species present, but could enable assembly of partial or complete genomes for the organisms identified [e.g. 48]. While such approaches are expensive and likely to be available only to a few at present, the holistic information on the eukaryome of NHPs likely to be obtained by such approaches makes them very attractive and we look forward to seeing the data emerge in the next few years.

Concluding Remarks

Currently, at least six *Entamoeba* species with valid published names have been confirmed by molecular analysis in NHPs: *E. coli, E. polecki, E. histolytica, E. nuttalli, E. dispar* and *E. hartmanni*. However, in addition there are multiple subtypes within *E. coli* and *E. polecki*, plus organisms with no name but distinct gene sequences (*Entamoeba* RL3, RL6, RL7 and CL8). This remarkable expansion in known diversity has been driven largely by the use of molecular techniques that have facilitated the identification of many novel and previously unrecognised *Entamoeba* species in NHPs.

However, many points remain to be clarified (see "Outstanding Questions"). It is unclear whether *E. moshkovskii, E. bangladeshi* and *E. gingivalis* colonise NHPs as well as humans.

373 Novel sequences with no linked species name are likely to continue to be detected in NHPs
374 around the world. This search for new types of *Entamoeba* in NHPs is essential as it remains
375 to be proven whether only *E. nuttalli* is responsible for morbidity and mortality in these hosts.
376 However, unless the correct approaches are used, such organisms will remain undiscovered.
377
378 We now know that NHPs are infected by both NHP-restricted and human-infective *Entamoeba*
379 species. Morphological diagnosis of *Entamoeba* species will always be problematic, but most
380 molecular approaches used to date may also be considerably underestimating the prevalence,
381 diversity, and distribution of *Entamoeba* in NHPs. At the same time, insufficient taxon
382 sampling and the heavy focus on humans may well have led us to inaccurate conclusions about
383 *Entamoeba* evolution. Fortunately, interest in the eukaryotic microbiome is growing in parallel
384 with improvements in technology, and it is likely that within the next few years a better
385 understanding of the evolution and host ranges of *Entamoeba* in NHPs will emerge.
386
387 Metagenomic analyses could allow the use of genes other than SSU-rDNA for phylogenetic
388 analyses. Obtaining sequence data for other genes is difficult - if not impossible - using
389 traditional molecular approaches and DNA from faecal samples. Multigene phylogenies may
390 well provide greater resolution that could confirm or refute our current views of relationships
391 within *Entamoeba*. Greater resolution is essential for evaluating the relative importance of
392 cospeciation and host-switching in the evolution of primate *Entamoeba* species. It seems likely
393 that these data will start to become available in the near future.
394
395 A recent study showed a significant reduction in the gut microbiome diversity of captive NHPs,
396 with a shift occurring from a wild NHP microbiome state toward a modern human microbiome
397 state [49]. Whether alterations in the lifestyle and diet of captive NHPs or the disruption of

398 normal hierarchical social behavior [50] has led to this perturbation of their gut microbiome,
399 the change may predispose captive NHPs to infection with certain *Entamoeba* spp, normally
400 confined to humans. Comparison of gut microbiomes across NHPs living in the wild,
401 semicaptivity and captivity using sequencing of both bacteria and *Entamoeba* SSU rDNA, is
402 already possible. Such data will allow us to investigate the correlation between microbiota
403 signatures and prevalence of specific *Entamoeba* species in NHPs.
404
405 There is much more to learn regarding both the microbiome and the eukaryome of NHPs,
406 especially those in the wild. There has been a strong focus on Old World primates, in particular
407 macaques, while New World primates are significantly underrepresented and prosimians have
408 not been studied. It is hoped that the range of species sampled will broaden, otherwise we will
409 continue to have a rather limited view of *Entamoeba* diversity in NHPs.
410
411 **Acknowledgments**
412 We would like to thank all the scientists whose research discoveries have advanced our
413 knowledge of the epidemiology of *Entamoeba* in non-human primates and enabled us to write
414 this review.

References

1. Lozano, R. *et al.* (2012) Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 380, 2095–2128

2. Beaver, P.C. *et al.* (1988) Invasive amebiasis in naturally infected New World and Old World monkeys with and without clinical disease. *Am. J. Trop. Med. Hyg.* 39, 343–352

3. Dobell, C. (1919) *The amoebae living in Man. A zoological monograph*. J. Bale, Sons, and Danielson, London, UK

4. Dobell, C. (1928) Researches on the intestinal protozoa of monkeys and man. II Description of the whole life-history of *Entamoeba histolytica* in cultures. *Parasitology* 20, 365–412

5. Burrows, R.B. (1957) *Endamoeba hartmanni*. *Am. J. Hyg.* 65, 172–188

6. Brumpt, E. (1925) Étude sommaire de *l'Entamoeba dispar* n. sp. Amibe à kystes quadrinucléés, parasite de l'homme. Bull. *Acad. Méd.* (Paris) 94, 943–952

7. Brumpt, E. (1928) Differentiation of human intestinal amoebae with four-nucleated cysts. *Trans. R. Soc. Trop. Med. Hyg.* 22, 101–114 (Discussion on pp. 115–124)

439 8. Martínez-Palomo A. *et al.* (1973) Selective agglutination of pathogenic strains of
440 *Entamoeba histolytica* induced by con A. *Nat. New Biol.* 245, 186–187
441
442 9. Sargeaunt, P.G. *et al.* (1978) The differentiation of invasive and non-invasive
443 *Entamoeba histolytica* by isoenzyme electrophoresis. *Trans. R. Soc. Trop. Med. Hyg.* 72, 519–
444 521
445
446 10. Diamond, L.S. and Clark, C.G. (1993) A redescription of *Entamoeba histolytica*
447 Schaudinn, 1903 (emended Walker, 1911) separating it from *Entamoeba dispar* Brumpt, 1925.
448 *J. Eukaryot. Microbiol.* 40, 340–344
449
450 11. Clark, C.G. and Diamond, L.S. (1991) The Laredo strain and other *Entamoeba*
451 *histolytic*a-like amoebae are *Entamoeba moshkovskii*. *Mol. Biochem. Parasitol.* 46, 11–18
452
453 12. Verweij, J.J. *et al.* (2001) Genetic variation among human isolates of uninucleated cyst-
454 producing *Entamoeba* species. *J. Clin. Microbiol.* 39, 1644–1646
455
456 13. Royer, T.L. *et al.* (2012) *Entamoeba bangladeshi* nov. sp., Bangladesh. *Emerg. Infect.*
457 *Dis.* 18, 1543–1545
458
459 14. Stensvold, C.R. *et al.* (2011) Increased sampling reveals novel lineages of *Entamoeba*:
460 consequences of genetic diversity and host specificity for taxonomy and molecular detection.
461 *Protist* 162, 525–541
462
463 15. Pritt, B.S. and Clark, C.G. (2008) Amebiasis. *Mayo Clin. Proc.* 83, 1154–1160

16. Haq, A. *et al.* (1985) Experimental infection of rhesus monkeys with *Entamoeba histolytica* mimics human infection. *Lab. Anim. Sci.* 35, 481–484

17. Tachibana, H. *et al.* (2007) An *Entamoeba* sp. strain isolated from rhesus monkey is virulent but genetically different from *Entamoeba histolytica*. *Mol. Biochem. Parasitol.* 153, 107–114

18. Tachibana, H. *et al.* (2009) Isolation and characterization of a potentially virulent species *Entamoeba nuttalli* from captive Japanese macaques. *Parasitology* 136, 1169–1177

19. Suzuki, J. *et al.* (2007) Profiles of a pathogenic *Entamoeba histolytica*-like variant with variations in the nucleotide sequence of the small subunit ribosomal RNA isolated from a primate (De Brazza's guenon). *J. Zoo Wildl. Med.* 38, 471–474

20. Takano, J. *et al.* (2009) DNA characterization of simian *Entamoeba histolytica*-like strains to differentiate them from *Entamoeba histolytica*. *Parasitol. Res.* 10, 929–937

21. Levecke, B. *et al.* (2010) Molecular identification of *Entamoeba* spp. in captive nonhuman primates. *J. Clin. Microbiol.* 48, 2988–2990

22. Castellani, A. (1908) Note on a liver abscess of amoebic origin in a monkey. *Parasitology* 1, 101–102

488 23. Dobell, C. (1931) Researches on the intestinal protozoa of monkeys and man. IV An
489 experimental study of the *histolytica*-like species of *Entamoeba* living naturally in macaques
490 in cultures. *Parasitology* 23, 1–72
491
492 24. Tachibana, H. *et al.* (2016) Isolation and molecular characterization of *Entamoeba*
493 *nuttalli* strains showing novel isoenzyme patterns from wild Toque Macaques in Sri Lanka. *J.*
494 *Eukaryot. Microbiol.* 6, 171–180
495
496 25. Suzuki, J. *et al.* (2008) A survey of amoebic infections and differentiation of an
497 *Entamoeba histolytica*-like variant (JSK2004) in nonhuman primates by a multiplex
498 polymerase chain reaction. *J. Zoo Wildl. Med.* 39, 370–379
499
500 26. Tachibana, H. *et al.* (2013) Prevalence of *Entamoeba nuttalli* infection in wild rhesus
501 macaques in Nepal and characterization of the parasite isolates. *Parasitol. Int.* 62, 230–235
502
503 27. Levecke, B. *et al.* (2015) Transmission of *Entamoeba nuttalli* and *Trichuris trichiura*
504 from Nonhuman Primates to Humans. *Emerg. Infect. Dis.* 21, 1871–1872
505
506 28. Sargeaunt, P.G. (1987) The reliability of *Entamoeba histolytica* zymodemes in clinical
507 diagnosis. *Parasitol. Today* 3, 40–43
508
509 29. Abd Alla, M.D. *et al.* (2012) Efficacy of a Gal-lectin subunit vaccine against
510 experimental *Entamoeba histolytica* infection and colitis in baboons (*Papio* sp.). *Vaccine* 30,
511 3068–3075
512

30. Rivera, W.L. *et al.* (2010) *Entamoeba histolytica* and *E. dispar* infections in captive macaques (*Macaca fascicularis*) in the Philippines. *Primates* 51, 69–74

31. Silberman, J.D. *et al.* (1999) Phylogeny of the genera *Entamoeba* and *Endolimax* as deduced from small subunit ribosomal RNA gene sequence analysis. *Mol. Biol. Evol.* 16, 1740–1751

32. Burrows, R.B. (1959) Morphological differentiation of *Entamoeba hartmanni* and *E. polecki* from *E. histolytica*. *Am. J. Trop. Med. Hyg.* 8, 583–589

33. Ponce Gordo, F. *et al.* (2004) *Entamoeba struthionis* n.sp. (Sarcomastigophora: Endamoebidae) from ostriches (*Struthio camelus*). *Vet. Parasitol.* 119, 327–335

34. Jacob, A.S. *et al.* (2016) Expanding the *Entamoeba* universe: new hosts yield novel ribosomal lineages. *J. Eukaryot. Microbiol.* 63, 69–78

35. Tuda, J. *et al.* (2016) 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.* 63, 572–577

36. Nolan, M.J. *et al.* (2017) Molecular characterisation of protist parasites in human-habituated mountain gorillas (*Gorilla beringei beringei*), humans and livestock, from Bwindi impenetrable National Park, Uganda. *Parasit. Vectors* 10, 340

537 37. Matthews, J.R. (1919) A mensurative study of the cysts of *Entamoeba coli*. *Ann. Trop.*
538 *Med. Parasitol.* 12, 259–272
539
540 38. Dobell, C. (1936) Researches on the intestinal protozoa of monkeys and man. VIII. An
541 experimental study of some simian strains of "*Entamoeba coli*". *Parasitology* 28, 541–593
542
543 39. Villanueva-García, C. *et al.* (2017). New *Entamoeba* group in howler monkeys
544 (*Alouatta* spp.) associated with parasites of reptiles. *Parasitol. Res.* 116, 2341–2346
545
546 40. Clark, C.G. and Diamond, L.S. (1997) Intraspecific variation and phylogenetic
547 relationships in the genus *Entamoeba* as revealed by riboprinting. *J. Eukaryot. Microbiol.* 44,
548 142–154
549
550 41. Al-Areeqi, M.A. *et al.* (2017) First molecular epidemiology of *Entamoeba histolytica*,
551 *E. dispar* and *E. moshkovskii* infections in Yemen: different species-specific associated risk
552 factors. *Trop. Med. Int. Health* 22, 493–504
553
554 42. López, M.C. *et al.* (2015) Molecular epidemiology of *Entamoeba*: first description of
555 *Entamoeba moshkovskii* in a rural area from central Colombia. *PLoS One* 10, e0140302
556
557 43. Nath, J. *et al.* (2015) Molecular epidemiology of amoebiasis: a cross-sectional study
558 among North East Indian population. *PLoS Negl. Trop. Dis.* 9, e0004225
559
560 44. Garcia, G. *et al*. (2014) Molecular epidemiology and genetic diversity of *Entamoeba*
561 species in a chelonian collection. *J. Med. Microbiol.* 63, 271–283

562
563
564 45. Levine, N.D. (1973) *Protozoan parasites of domestic animals and of man. Second*
565 *Edition*. Burgess Pub. Co., Minneapolis, MN, USA
566
567 46. Jirků-Pomajbíková, K. *et al.* (2016) Molecular identification of *Entamoeba* species in
568 savanna woodland chimpanzees (*Pan troglodytes schweinfurthii*). *Parasitology* 143, 741–748
569
570 47. Wegener Parfrey, L. *et al.* (2014) Communities of microbial eukaryotes in the
571 mammalian gut within the context of environmental eukaryotic diversity. *Front. Microbiol.* 5,
572 298
573
574 48. Beghini, F. *et al.* (2017) Large-scale comparative metagenomics of *Blastocystis*, a
575 common member of the human gut microbiome. *ISME J.* 11, 2848–2863
576
577 49. Clayton, J.B. *et al*. (2016) Captivity humanizes the primate microbiome. *Proc, Natl,*
578 *Acad, Sci, USA* 113, 10376–10381
579
580 50. Perofsky, A.C. *et al*. (2017) Hierarchical social networks shape gut microbial
581 composition in wild Verreaux's sifaka. *Proc. R. Soc. B* 284, 20172274
582
583 51. Feng, M. *et al.* (2013) Prevalence and genetic diversity of *Entamoeba* species infecting
584 macaques in southwest China. *Parasitol. Res.* 112, 1529–1536
585

586 52. Guan, Y. *et al.* (2016) Comparative analysis of genotypic diversity in *Entamoeba*
587 *nuttalli* isolates from Tibetan macaques and rhesus macaques in China. *Infect. Genet. Evol.* 38.
588 126–131
589
590 53. Debenham, J.J. *et al.* (2017) Occurrence of *Giardia*, *Cryptosporidium*, and *Entamoeba*
591 in wild rhesus macaques (*Macaca mulatta*) living in urban and semi-rural North-West India.
592 *Int. J. Parasitol. Parasites Wildl.* 6, 29–34
593
594 54. Dong, H. *et al.* (2017) Prevalence, molecular epidemiology, and zoonotic potential of
595 *Entamoeba* spp. in nonhuman primates in China. *Infect. Genet. Evol.* 54, 216–220
596
597 55. Al-Mayali, H.M.H. and Al-Abodi, H.R.J. (2017) Molecular characterization of
598 *Entamoeba* spp. in Al-Qadisiya province, Iraq. *Al-Kufa Univ. J. Biol*. 9, 374–386
599
600 56. Ragazzo, L.J. *et al.* (2018) *Entamoeba histolytica* infection in wild lemurs associated
601 with proximity to humans. *Vet. Parasitol.* 249, 98–101
602

Glossary

Conditional lineage (CL): an *Entamoeba* identified as likely to be distinct based on sequencing of partial SSU-rDNA, but for which sufficient data are not yet available. See RL, below.

Isoenzymes: each of two or more sequence variants of an enzyme that exhibit different migration in electrophoresis gels due to charge differences.

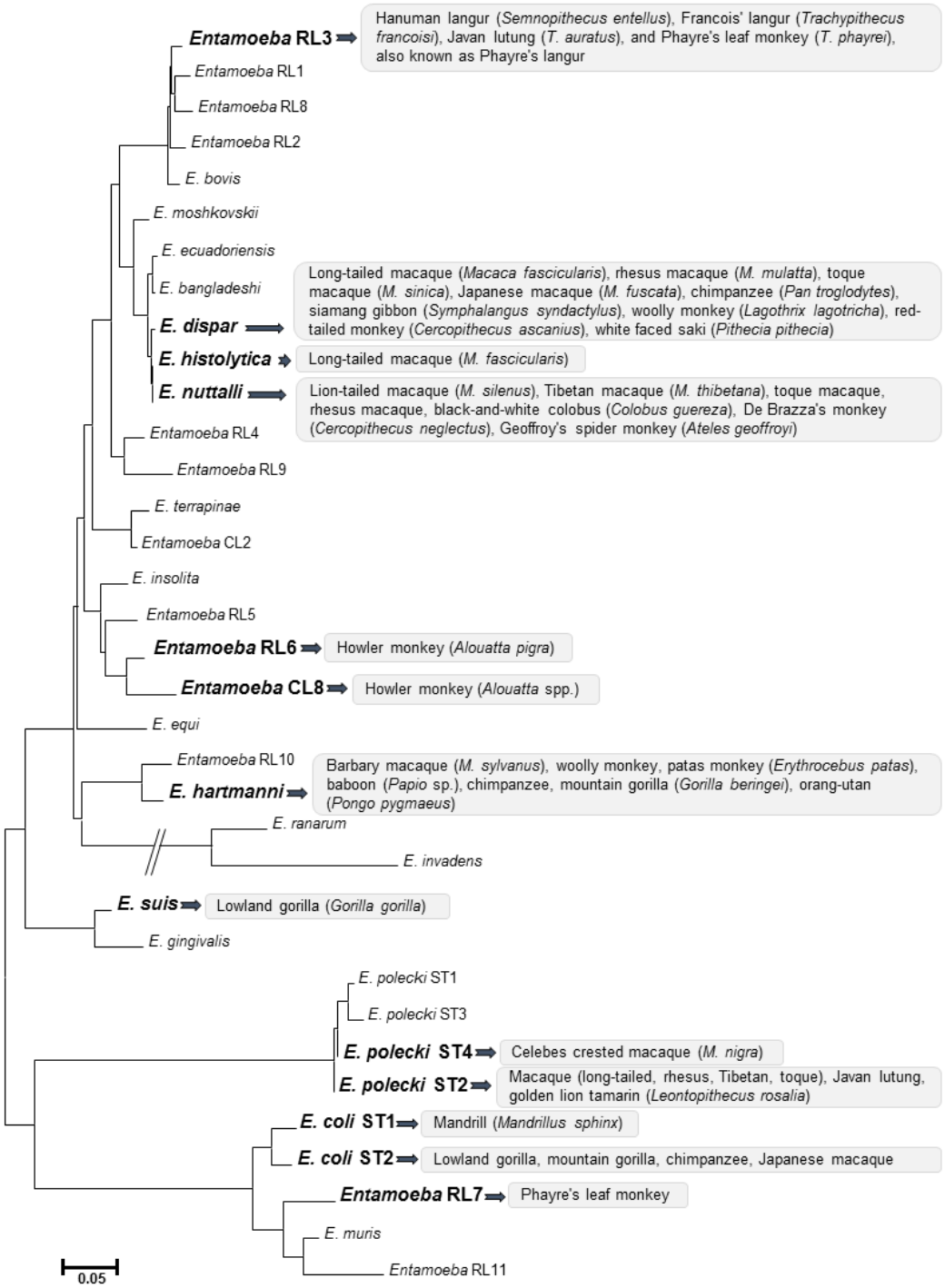
Non-human primates (NHPs): all members of the order Primates other than humans; NHPs share many similarities with humans in terms of physiology, anatomy, immunology, and neurology, but are very diverse in their ecology, diet, etc. The split between humans and NHPs is an artificial one, as humans are much more closely related to some NHPs than others.

Ribosomal lineage (RL): an *Entamoeba* taxon identified as distinct by sequencing of its complete SSU-rDNA gene. Often no corresponding morphological data are available. In other groups of organisms these are often called operational taxonomic units (OTUs) but in this case, it is clear that they belong to the genus *Entamoeba*.

Small subunit ribosomal RNA gene (SSU-rDNA): the gene encoding the smaller of the two major RNA components of the ribosome, also known as 18S rDNA. This gene is the most widely used single locus for phylogenetic analyses in eukaryotes and bacteria. In *Entamoeba,* the gene size generally falls between 1800 and 2200 bases.

Subtype: a discrete genetic clade within a named species.

628 **Figure 1: Phylogenetic relationships among *Entamoeba* species.** The phylogenetic tree
629 shown is modified from Figure 1 in Jacob et al. [34]. Names in bold lettering are those that
630 have been identified by sequencing of SSU-rDNA in NHPs. Adjacent to the *Entamoeba* names
631 are those of the NHP species (wild or captive) in which the *Entamoeba* has been identified.
632



633

Dobell nomenclature[a]	Current species names	Identified in primates (incl. humans)	Molecular identification in NHPs
E. histolytica	*E. histolytica*	*E. histolytica*	Y[b]
	E. dispar	*E. dispar*	Y
	E. hartmanni	*E. hartmanni*	Y
	E. nuttalli	*E. nuttalli*	Y
	E. moshkovskii	*E. moshkovskii* (complex)	N
	E. polecki	*E. polecki* ST1 [c]	N
		E. polecki ST4	Y
	E. chattoni	*E. polecki* ST2	Y
	E. struthionis	*E. polecki* ST3	N
	E. bangladeshi	*E. bangladeshi*	N
	E. suis	*E. suis*	Y[b]
E. coli	*E. coli*	*E. coli* ST1	Y[b]
		E. coli ST2	Y
E. gingivalis	*E. gingivalis*	*E. gingivalis* ribodeme 1 [d]	N
		E. gingivalis ribodeme 2	N
	None	*Entamoeba* RL3 [e]	Y
		Entamoeba RL6	Y
		Entamoeba RL7	Y
		Entamoeba CL8 [f]	Y

Table 1. Correspondence between historic, binomial, and sequence-based nomenclature for *Entamoeba* species in primates.

[a] Dobell's nomenclature is that proposed in his 1919 monograph [3].

[b]Identified in captive NHPs only, to date.

[c] Subtypes (ST) are distinct small-subunit ribosomal DNA sequence variants that clearly fall within a named species.

[d] Ribodemes are small-subunit ribosomal DNA variants detected by restriction enzymes.

[e] Ribosomal (RL) [14]] lineages indicate complete small-subunit ribosomal DNA sequences that are clearly distinct from all named species.

[f] Conditional (CL) [34] lineages indicate partial small-subunit ribosomal DNA sequences that are clearly distinct from all named species.

NHP species	Type of amplification	Total no. of samples	Species identified (no. of samples)	Reference	Notes
Rhesus macaques *(Macaca mulatta*)	Species-specific	715	*E. nuttalli* (440), *E. dispar* (16), *E. coli* (574), *E. polecki* ST2 (649)	51	a
Rhesus macaque (*Macaca mulatta*)	Species-specific	112	*E. nuttalli* (57), *E. dispar* (13), *E. coli* (83), *E. polecki* ST2 (96)	26	b
Tibetan macaque (*Macaca thibetana*)	Species-specific	89	*E. nuttalli* (15), *E. coli* (37), *E. polecki* ST2 (59)	52	c
Savannah woodland chimpanzee (*Pan troglodytes schweinfurthii*)	Genus- and species-specific	107	*E. hartmanni* (32), *E. dispar* (10), *E. coli* ST2 (33)	46	d
Celebes crested macaque (*Macaca nigra*)	Species/subtype-specific	77	*E. polecki* ST4 (75)	35	e
Toque macaque (*Macaca sinica*)	Species-specific	227	*E. nuttalli* (42), *E. dispar* (1), *E. coli* (40), *E. polecki* ST2 (197)	24	f
Rhesus macaque (*Macaca mulatta*)	Genus- and species-specific	128	*E. coli* (63), unidentified *Entamoeba* (65)	53	g
Mountain gorilla (*Gorilla beringei beringei*)	Genus-specific	68	*E. coli* ST2 (4), *E. hartmanni* (33)	36	h
Howler monkeys (*Alouatta palliata* and *A. pigra*)	Genus-specific	155	*Entamoeba* CL8 (6 from *A. pigra*, 1 from *A. palliata*), *Entamoeba* RL6 (1 from *A. pigra*)	39	i

Table 2. Summary of results from molecular screening of faecal samples from wild NHP populations*

* The publication by Dong et al. [54] includes data on several NHP species in China (mostly *Macaca mulatta* and *M. fascicularis*) but it is not possible to identify which results came from sampling wild populations. Samples were tested by species-specific amplification for *E. histolytica,*

E. nuttalli, E. dispar, E. moshkovskii, E. coli, and *E. polecki* ST2. Only *E. coli* and *E. dispar* were detected. No tests for *E. hartmanni* or other *E. polecki* subtypes were performed.

a: Authors also tested captive macaques; these are excluded from the table. Tested for *E. histolytica*, *E. dispar*, *E. nuttalli*, *E. coli*, and *E. polecki* ST2 only. No test for *E. hartmanni*, *E. moshkovskii* or other *E. polecki* subtypes.

b: Tested for *E. histolytica*, *E. dispar*, *E. nuttalli*, *E. moshkovskii*, *E. coli*, and *E. polecki* ST2 only. No test for *E. hartmanni* or other *E. polecki* subtypes.

c: Tested for *E. histolytica*, *E. dispar*, *E. nuttalli*, *E. coli*, and *E. polecki* ST2 only. No test for *E. hartmanni*, *E. moshkovskii* or other *E. polecki* subtypes.

d: Genus-PCR-positive samples were tested for *E. histolytica*, *E. nuttalli*, *E. dispar*, *E. moshkovskii*, *E. coli*, and *E. polecki* ST2. Genus-PCR positive, but species-specific PCR negative samples were sequenced and identified as *E. hartmanni*.

e: Tested for *E. histolytica*, *E. dispar*, *E. nuttalli*, *E. moshkovskii*, *E. coli*, and *E. polecki* ST1, ST2 and ST4. No test for *E. hartmanni* or *E. polecki* ST3.

f: Tested for *E. histolytica*, *E. dispar*, *E. nuttalli*, *E. moshkovskii*, *E. coli*, and *E. polecki* ST2. No test for *E. hartmanni* or other *E. polecki* subtypes.

g: Genus-PCR positive samples were tested for *E. coli*. Multiplex PCR for *E. histolytica*, *E. dispar* and *E. moshkovskii* on all samples. No test for *E. nuttalli*, *E. polecki*, or *E. hartmanni*.

h: Sequencing of Genus-PCR positive amplicons identified only these two species.

i: Sequencing of Genus-PCR positive amplicons identified only these two organisms.