Jacob, AS; Busby, EJ; Levy, AD; Komm, N; Clark, CG (2015) Expanding the Entamoeba Universe: New Hosts Yield Novel Ribosomal Lineages. The Journal of eukaryotic microbiology, 63 (1). pp. 69-78. ISSN 1066-5234 DOI: https://doi.org/10.1111/jeu.12249

DOI: 10.1111/jeu.12249

Running head: Expanding the *Entamoeba* universe

Expanding the *Entamoeba* Universe: New Hosts Yield Novel Ribosomal Lineages [1]

Alison S. Jacob[a], Eloise J. Busby[a], Abigail D. Levy[a, 2], Natasha Komm[a] and C. Graham Clark[a]

a Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London WC1E 7HT, England, GB

Key words: *Entamoeba,* Genetic Diversity, Phylogeny, Ribosomal RNA, Sequencing

1 – This paper is dedicated to Govinda S. Visvesvara on the occasion of his retirement.
2 – Current Address: Histocompatibility & Immunogenetics Department, Hammersmith Hospital, Imperial College Healthcare NHS Trust, London W12 0HS, England, GB

Corresponding Author: C. Graham Clark, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, England, GB. Telephone number: +44-20-7927-2351; FAX number: +44-20-7636-8739; email: graham.clark@lshtm.ac.uk

ABSTRACT

Removing the requirement for cell culture has led to a substantial increase in the number of lineages of *Entamoeba* recognized as distinct. Surveying the range of potential host species for this parasite genus has barely been started and it is clear that additional sampling of the same host in different locations often identifies additional diversity. In this study, using small subunit ribosomal RNA gene sequencing, we identify four new lineages of *Entamoeba*, including the first report of *Entamoeba* from an elephant, and extend the host range of some previously described lineages. Additionally, examination of microbiome data from a number of host animals suggests that substantial *Entamoeba* diversity remains to be uncovered.

Keywords: Diversity; next generation sequencing; ribosomal RNA; phylogeny;

Over the past 25 years, our understanding of diversity in the genus *Entamoeba* has increased significantly as a result of complementary developments in DNA amplification, purification and sequencing. Traditionally, naming of species in *Entamoeba* was based on a mixture of host identity and parasite morphology, the latter being rather limited in these amoeboid organisms and the former being of debatable value due to uncertainty over host ranges of the parasites. DNA sequencing allows quantitative measurement of similarity that is not dependent on such characters and although it is not without its own limitations, it has fundamentally changed our approach to studying diversity in organisms such as *Entamoeba*.

In this time period, the study of *Entamoeba* has gone from being dependent on stable laboratory cultures of parasites, preferably in the axenic form, to DNA analysis of organisms directly from stool samples in the absence of even microscopic investigation. The latter aspect has been problematic as it is not possible to assign new sequences to previously named species where the original description is dependent on morphology. For this reason, many new and distinct *Entamoeba* sequences have been assigned to ‘ribosomal lineages’ rather than species to reflect the absence of morphological information (Stensvold et al. 2011).

The new approach is dependent on the reliability of DNA purification from stool samples, which are notorious for the presence of enzyme inhibitors, and the specificity of the primers used for PCR amplification. In addition, investigation of *Entamoeba* in such samples is largely limited to the ribosomal RNA genes due to the complexity of the DNAs extracted from stool, which often contains DNA from multiple other parasites and may include multiple *Entamoeba* species. The elimination of culture dependency has led to a dramatic expansion in the number of genetically distinct *Entamoeba* organisms being recognized but also to a greater understanding of sequence variability within species due to the relative ease with which multiple samples can be studied in parallel. The present report contributes information on several novel *Entamoeba* lineages as well as variation within and host range of known species.

MATERIALS AND METHODS

Cultures

E. bangladeshi strains 8111 and 8237 were received as xenic cultures from Dr Rashidul Haque, ICDDR, B, Bangladesh. Partial sequences of their small subunit ribosomal RNA genes (SSU rDNA) were published in the original species description (Royer et al. 2012) but as these only covered ca. 20% of the full gene, the complete sequences were obtained to allow more accurate phylogenetic investigation. The organisms were grown in LYSGM with 5% adult bovine serum and rice starch at 22 °C and subcultured twice weekly (Clark and Diamond 2002).

E. invadens VK-1 was received as an axenic culture from Dr Avelina Espinosa (Roger Williams University, United States) and grown in LYI-S-2 (Clark and Diamond 2002) with 15% adult bovine serum at 22 °C. The complete SSU rDNA was sequenced to investigate intra-specific variation in this species.

Stool samples

Stool samples came from a variety of sources. Most were collected by MSc students as part of a parasitology field trip in two consecutive years. Asian elephant (*Elephas maxima*) stool samples were obtained from Amsterdam Zoo, courtesy of Mark J. Hoyer and Daphne Valk through Dr Bruno Levecke, University of Ghent. Samples were either extracted as fresh material or stored (at 4 °C) in 70% ethanol at a ratio of 3:1 (v:v) for later processing. Some of the DNA samples used were from a previously published study (Alfellani et al. 2013).

DNA purification from stool

When present, ethanol was removed from stool samples by washing in phosphate buffered saline (pH = 8) three times prior to processing. DNA was extracted from stool using the Qiagen DNA stool minikit (Qiagen, Hilden, Germany) according to the manufacturer’s instructions.

Amplification and sequencing

Primers used, amplification and sequencing conditions were all essentially as described previously unless otherwise stated (Stensvold et al. 2011). Initial analysis of samples involved amplification of purified DNA using the genus-specific primer pairs Entam 1/2 (Verweij and Stensvold 2014) or Entagen F/R (Stensvold et al. 2011). PCR conditions consisted of 30 cycles of 30 sec each at 94 °C, 59 °C and 72 °C with a final extension of 72 °C for 10 min.

Phylogenetic analysis

New sequences were aligned with reference sequences from Stensvold et al. (2011), using the alignment tool MUSCLE as implemented in MEGA 5 (Tamura et al., 2011). The alignment was edited manually to remove regions of ambiguity resulting in an alignment of 1,447 positions. Phylogenetic analyses were performed using distance (Neighbor-Joining (NJ); distance matrix in Supplementary Data File S2) and Maximum Likelihood (ML) algorithms as implemented in MEGA5 and Bayesian analysis (MrBayes 3.1.5; Huelsenbeck and Ronquist, 2001). Bayesian and ML analysis used a General Time Reversible (GTR) model of nucleotide substitution with four categories of among-site rate variation and the proportion of invariant sites, the best model selected by ModelTest, implemented in MEGA5. Statistical support for distance and maximum likelihood trees was evaluated using bootstrapping (1,000 replicates). Bayesian analysis used four Markov chain Monte Carlo (MCMC) strands, 1,000,000 generations, with trees sampled every 100 generations. The resulting average standard deviation of split frequencies was less than 0.01. A consensus tree was produced after excluding an initial burn-in of 25% of the samples, as recommended.

For the analysis of the microbiome sequence data, sequences were extracted from the curated SILVA 108 database (http://qiime.org/home_static/dataFiles.html) generated by Parfrey et al. (2014) which are also available as Supplementary Data File S1. After alignment with relevant reference sequences and editing, as above, the same phylogenetic methods were employed, except that for Bayesian analysis 2,000,000 generations were used and the standard deviation of split frequencies stabilized at 0.027.

Terminology

Because non-standard nomenclature is used to describe *Entamoeba* diversity, we here define some of the terms used. These explain our working criteria based on the data available at this time.

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. We will use the term “species” 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 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 RL numbers. It is possible that RLs could become species in the future if morphological and other relevant data become available, but we do not think it 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, we propose identifying it as a conditional lineage. Such lineages are likely to represent novel organisms and to be “promoted” to RLs or species when more data become available.

RESULTS AND DISCUSSION

Molecular species identification and survey overview

In addition to a few cultures, stool samples were obtained from a wide range of hosts in a number of locations, including farms and zoos, as well as from trapped wild hosts.

Culture material gave no problems with either amplification or sequencing. However, with stool DNA, *Entamoeba* sequences were preferentially amplified when present but, when absent, non-specific products from the SSU rDNA of fungi and plants were sometimes produced. It is likely that some of the unreadable sequences obtained were due to mixed products of multiple origins. PCR positivity varied dramatically among hosts, with sheep and cattle being the hosts most frequently infected with *Entamoeba*.

Where a partial sequence obtained showed 98% or greater identity to reference sequences, the *Entamoeba* species or RL was considered to have been identified and no further investigation was undertaken. However samples yielding sequences with lower identities, 95% or less, were subjected to further amplification and sequencing, usually involving broad-specificity primers paired with sequence-specific primers designed from the initial sequence data obtained. DNA sample amount was a significant limitation, in some cases leading to the DNA being depleted before the complete gene sequence could be obtained. There were nevertheless sufficient data to allow phylogenetic analyses to be undertaken in most cases.

The results from our survey implementing this approach are presented in Table 1 and include the identification of four new ribosomal lineages (RLs 8, 9, 10 and 11). Phylogenetic analysis using their SSU rDNA sequences (Fig. 1) identified their relationships to previously described species and lineages.

***Entamoeba* from artiodactyls**

As the majority of ungulate samples in our study were obtained from ruminant artiodactyl species it is of little surprise that this group displays the highest number of *Entamoeba* PCR positive samples, given previous results (Stensvold et al. 2011). We did note however, that the amplification of non-specific PCR products (fungi, plants and other protists) was a particular problem when dealing with ruminant faecal material and that this may have precluded the identification of *Entamoeba* in some samples.

Entamoeba *in cattle.* In cattle, an *Entamoeba* positivity rate of 44.4% was observed. *E. bovis* was the predominant species with single occurrences of *Entamoeba* RL2, *Entamoeba* RL4 and *E. moshkovskii*. Three samples were found to contain mixed infections. The most abundant species, as determined from the highest peaks in the sequence trace chromatograms, was *E. bovis* and the minority species was undetermined. This observation is in keeping with our previous findings that it is possible for a single animal to be infected with more than one lineage of *Entamoeba* (Stensvold et al. 2011).

It is of particular interest that *E. moshkovskii* was detected in cattle from Devon, UK. To our knowledge this represents the first example of this species being detected in a ruminant animal, having been reported previously only in humans (Heredia et al. 2012) and aquatic turtles (Garcia et al. 2014), in addition to environmental samples. The detection of *Entamoeba* RL4 in Hertfordshire, UK, further expands the geographic distribution of this lineage, which has been detected previously only in cattle from Libya and Estonia (Stensvold et al. 2011).

A new lineage was discovered in cattle from Devon (sequence Cow5) which we identify here as *Entamoeba* RL8 (Fig. 1). Our phylogenetic analysis suggests that *Entamoeba* RL8 is most closely related to *Entamoeba* RL1, a sequence from Roe Deer, also a ruminant artiodactyl, albeit with weak support.

Entamoeba *in sheep.* The prevalence of *Entamoeba* in sheep (49.1%) was slightly higher than that found in cattle. Again, *E. bovis* was the predominate species and the remainder of the positive samples were represented by the single detection of *Entamoeba* RL2 and three mixed *E. bovis* samples. The identification of *Entamoeba* RL2 in this group further expands the host range of this lineage, which has previously only been found in cattle (Stensvold et al. 2011).

To date, the molecular sampling of sheep populations has failed to identify a distinct lineage that could correspond to the species *E. ovis* (Noble and Noble, 1952). Phylogenetic analysis of two sheep derived DNA sequences (Stensvold et al. 2010) and the sequence data from the present survey (derived from 27 *Entamoeba*-positive specimens) show that sequences obtained from sheep do not form a unique host-specific clade. Based on this evidence, we therefore suggest that *E. ovis* is a junior synonym and that *E. bovis* infects both cattle and sheep.

Entamoeba *in deer.* Nine deer were sampled, with a positivity rate of 22.2%. The discovery of *E. bovis* in two fallow deer from Mauritius gives a new host for this species, having only been detected in reindeer (*Rangifer tarandus*) previously (Clark et al. 2006). Our data confirm that *E. bovis* has a wide geographic distribution and a wide host range which now encompasses cattle, sheep and deer, all of which are ruminant artiodactyls.

Entamoeba *in pigs.* In non-ruminant artiodactyls, without exception the species present in PCR positive samples (52.4%) derived from pigs was *E. polecki* ST1. Three specimens were found to have a mixed infection consisting of two closely related subtypes (ST1 and ST3). It is noteworthy that *E. suis* was not detected. A study of five pigs taken from a population of 148 displaying hemorrhagic colitis symptoms in Japan suggested that *E. suis* was responsible for the disease (Matsubayashi et al., 2014). Conversely, *E. suis* was detected in 28 specimens taken from pigs in Vietnam and no ill health was reported (Clark et al. 2006).

Entamoeba bovis-*related lineages.* The shaded area in Fig. 1 identifies the current phylogenetic relationships of *E. bovis* and related ribosomal lineages. We find that *E. bovis* and *Entamoeba* RLs 1–3 plus the newly defined *Entamoeba* RL8 form a strongly supported monophyletic clade. However, within this clade we were unable to resolve the relationships among the lineages. The grouping of *E. bovis* and *Entamoeba* RLs 1–3 together as a robust monophyletic clade with high bootstrap support and low resolution is consistent with previous tree reconstructions (Stensvold et al. 2011).

In need of further comment is the continued, seemingly incongruous, inclusion of *Entamoeba* RL3, a sequence isolated from langurs, within a clade otherwise consisting of sequences from ruminant artiodactyls. Langurs of the genera *Trachypithecus* and *Semnopithecus* are unique amongst primates for their possession of a ruminant-like stomach (Bauchop and Martucci 1968). We speculate that the similarity of the langur ruminant stomach to that found in artiodactyls provides environmental and physiological conditions in which an *Entamoeba* lineage related to *E. bovis* was able to colonize and subsequently become host-adapted. It should be pointed out that, at present, no stool samples of artiodactyls from East, South and Southeast Asia have been investigated.

Entamoeba *in microbiome data.* During the course of our survey work, a study was published that utilised high-throughput 454-amplicon pyrosequencing to systematically investigate the eukaryotic communities in mammalian gut microbiota, including *Entamoeba* species (Parfrey et al. 2014). The study included samples obtained from captive herbivores. Some of these were positive for *Entamoeba* and are listed in Table 2. A preliminary analysis of the sequences from this dataset revealed that one of the sequences obtained from an Okapi was identical to our newly identified lineage *Entamoeba* RL8, isolated from cattle. This serendipitous discovery prompted us to further analyze *Entamoeba* sequences isolated from this dataset and the phylogenetic reconstruction is presented in Fig. 2.

Since the data are derived from the short reads (ca 500 bp) generated by 454-amplicon pyrosequencing we are unable to assign new ribosomal lineages on the basis of the guidelines for *Entamoeba* nomenclature (Stensvold et al. 2011). Instead, we refer to the several potentially new lineages as “conditional lineages” (CL; see Methods section). The resultant phylogenetic tree shows the relationships of the 454 sequences related to *E. bovis* plus RLs 1-4 plus 8. Five newly defined *Entamoeba* clades, CLs3-7, are identified.

The shaded region of the resulting cladogram shows an expanded version of the *E. bovis* clade, which loosely comprises six internal clusters (Fig. 2). While the monophyly of the *E. bovis* clade is strongly supported in the ML analysis (95%) it is poorly supported in the distance-based analysis (67%), and is absent in the Bayesian analysis. The latter results from the inclusion of a single sequence from an Okapi (Okapi2 92564 in Fig. 2) within the *E. bovis* clade in this analysis only. Otherwise, the clade contains the same sequences in all analyses and has strong Bayesian posterior probability support. This sequence appears as a distinct lineage, identified as *Entamoeba* CL3, in both ML and NJ analyses.

A number of sequences were identified in two Okapi, a ruminant artiodactyl, sampled in this survey and both specimens revealed known and novel lineages of *Entamoeba*. A total of eight sequences were obtained from the Okapi2 sample and these included five belonging to the *E. bovis* clade. One corresponds to the newly defined *Entamoeba* RL8, robustly supported in all three analysis methods, while the remaining two sequences from this host (Okapi2 92564 and Okapi2 39254) were assigned to the putative ribosomal clades *Entamoeba* CL3, which has no close relatives, and *Entamoeba* CL4, respectively. The existence of *Entamoeba* CL4 is supported in ML (92%) and Bayesian inference, but not NJ analysis. In contrast, just three *Entamoeba* sequences were obtained from sample Okapi1: two of these fall into *Entamoeba* CL4 and one (Okapi16235) into the newly defined *Entamoeba* CL5 with high support in both ML and Bayesian inference analysis.

One of the sampled gazelles (Gazelle3) yielded an astonishing 14 distinct *Entamoeba* sequences, 11 of which cluster within the *E. bovis* clade, two with *Entamoeba* RL8 and one with *Entamoeba* RL1, which expands the host range of that lineage. A big horn sheep sample (BigHornSD) yielded 10 distinct sequences of which 7 were found to cluster within *E. bovis*, one with *Entamoeba* RL1 (also expanding the host range of this lineage), and two within the newly defined *Entamoeba* CL5.

There were three sequences found to have no close relatives in our phylogenetic reconstruction; sequences from a kangaroo (*Entamoeba* CL7), an okapi (*Entamoeba* CL3) and a wild ass (*Entamoeba* CL6). This survey also demonstrates that the same host species at the same location can carry different lineages of *Entamoeba*, in that the sequences obtained from two okapi living in the same herd showed differences in lineage representation.

Finally, the detection of *E. bovis* in a kangaroo represents a very different new host for this species. Macropods (kangaroos and wallabies) are foregut fermenters but do not have the ruminant stomach structure seen in cattle. It is possible that this *Entamoeba* was transient and not established within the host. Only further sampling of macropods will help to establish whether *E. bovis* is a normal member of their gut fauna.

The data presented in Fig. 2 reveal not only a remarkable degree of diversity within the known *E. bovis* and related lineages but gives an insight into *Entamoeba* diversity within a single host. Application of this approach to analyzing the eukaryotic microbiome in a range of host samples is likely to become the method of choice in the future for detecting diversity, although the short sequences obtained are not ideal for phylogenetic analyses.

***Entamoeba* from elephants**.

There have been three parasitological surveys of wild elephants in the past decade. The first documented nematode and ciliate populations in the stool of African forest elephants (Kinsella et al. 2004), a second reported helminth and coccidian parasites from African elephants in Botswana (Baines et al. 2015), while the other selectively concentrated on nematode eggs in Asian elephants (Hing et al. 2013). None of the studies reported finding *Entamoeba*.

To the best of our knowledge, this is the first report of *Entamoeba* in elephants. The sequence obtained from an Asian elephant living in Amsterdam Zoo represents a novel lineage, which we define as *Entamoeba* RL10. *E. moshkovskii* has also been identified, in an African elephant in a zoo setting (Table 2) (Parfrey et al. 2014). How widespread *Entamoeba* RL10 is, whether it is found in both Asian and African elephants, and whether it is specific to elephants will require additional surveys to be undertaken. It is also important that wild elephants be sampled to rule out a captivity-acquired infection. It is important that such surveys be molecular in nature, as no cysts were detected in the *Entamoeba*-positive sample from the Asian elephant. These might have been expected to be 4-nucleated owing to their relationship of RL10 to *E. hartmanni*. It could be speculated that the absence of cysts might explain why no *Entamoeba* has previously been reported from elephants (Kinsella et al. 2004) but raises the question of how it is transmitted.

Our phylogenetic analysis (Fig. 1) demonstrates that *Entamoeba* RL10 is closely related to *E. hartmanni* (having a pairwise sequence identity of 95.7%) and these two sequences consistently form a clade with very high bootstrap support in the recovered trees. The placement of *Entamoeba* RL10 as sister taxon to *E. hartmanni* is significant because the latter species has not been found to have any close relatives in previous phylogenetic reconstructions (Stensvold et al. 2011).

There is a tendency for NJ analyses to cluster the *Entamoeba* RL10/*E. hartmanni* clade with the *E. ranarum*/*E. invadens* clade. Support is weak but this finding is in agreement with previous studies (Stensvold et al. 2010, Clark et al. 2006). In contrast, our most recent study did not recover this relationship (Stensvold et al. 2011) suggesting that it may be sensitive to the sequences included in the alignment. It is clear that finding further lineages related to *E. hartmanni* will help to resolve the relationships of this well-defined clade.

Entamoeba moshkovskii and its relatives

The finding of *E. moshkovskii* in elephant and cattle further broadens the host range of this species. It is clear that this usually free-living species, which exists as a species complex consisting of multiple variants (Clark and Diamond 1997), is being detected in animal hosts more frequently, particularly in humans. Since the advent of molecular detection tools there has been growing interest in *E. moshkovskii* as there is speculation that it may be a facultative parasite (Heredia et al. 2012). Previous phylogenetic reconstructions have consistently placed *E. moshkovskii* as a sister lineage to a clade consisting of *E. dispar, E. nuttalli, E. histolytica* and *E. ecuadoriensis.* The latter has only been isolated once, from sewage, and like *E. moshkovskii* is considered to be potentially free-living (Stensvold et al. 2011, Stensvold et al. 2010). *E. bangladeshi* is the most recent species to be described from humans. In phylogenetic reconstructions based on partial SSU rDNA sequences, it was found to branch between *E. moshkovskii* and *E. ecuadoriensis* (Royer et al. 2012). However, our analysis based on the complete SSU rDNA sequences (KR025411 and KR025412) specifically positions *E. bangladeshi* in a highly supported clade with the latter species. *E. bangladeshi*, like *E. moshkovskii* and *E. ecuadoriensis*¸ is able to grow at both body temperature and room temperature, suggesting it might also be found in the environment in the future.

Entamoeba from horses

The species name *E. equi* was first used to describe 4-nucleated *Entamoeba* cysts in horses from South Africa (Fantham, 1921). It was subsequently resurrected for a DNA sequence obtained from a horse in Aberystwyth, UK (Clark et al. 2006). As the species name suggests, the host range of this species appears to be confined to the family Equidae, which also includes donkeys and zebras. The later finding of *E. equi* in a zebra (*Equus zebra hartmannae*) from a zoo in the UK (Stensvold et al. 2011) appeared to support this host range and species designation, although no cysts were seen in either sample.

A new ribosomal lineage, *Entamoeba* RL9, was detected in three horses from various locations in Devon. It occupies a position in the phylogenetic tree well removed from the distinct *E. equi* lineage (Fig. 1). Our phylogenetic reconstruction shows the placement of *Entamoeba* RL9 as a sister taxon to *Entamoeba* RL4 (a lineage associated with cattle) but this relationship is not well supported by either NJ or ML analyses. The exact position of *Entamoeba* RL9 may become clearer with further sampling. An intriguing observation is that although the multiple *Entamoeba* RL9 DNA sequences obtained are closely related, they are not identical.

In common with the findings for *E. equi*, cysts were not detectable microscopically in the new equine samples containing *Entamoeba* RL9 using standard methodologies. This suggests that horse physiology may be responsible for the absence of cysts rather than this being a species-specific trait of *E. equi*. It is also possible that the apparent absence of cysts in horses (and elephants) is related to their scarcity or to periodic shedding. Only more detailed investigation can solve this conundrum.

454-amplicon pyrosequencing data (Table 2) also identified *E. hartmanni* in a zebra (Parfrey et al. 2014), a new host for this species which has previously only been detected in primates. The latter observation again raises the question of whether zoo hosts are true natural reservoirs of the *Entamoeba* lineages being detected or if we are observing a captivity-acquired infection, in the same way that wild birds have been identified as a possible source of *Plasmodium* spp. in captive penguins (Leclerc et al. 2014). To date, information on *Entamoeba* in wild animals is sparse.

Entamoeba from rodents

There have been a number of morphological descriptions of *Entamoeba* spp. from rodents in the literature. These reports range from the morphologically indistinguishable 4-nucleated cysts of *E. histolytica* and *E. dispar* in rats (Mishra and Gonzalez 1975, Shafiyyah et al. 2012, Neal 1948) to the 8-nucleated cyst former *E. muris,* detectable in both wild (El-Ridi et al. 1987, Nateghpour et al. 2015) and laboratory rat populations (Won et al. 2006). Other 8-nucleated cyst forming species such as *E. funambulae* (Ray and Banik 1964) and *E. citelli* (Davis 1969, Diakou et al. 2015) have been reported in squirrel populations. Finally, there have been reports of *Entamoeba* spp. which also formed 8-nucleated cysts in Syrian hamsters (Neal 1947).

In contrast, there is a dearth of molecular sequence data relating to *Entamoeba* in rodents. Only recently has PCR been used to detect and differentiate *E. histolytica* and *E. dispar* in rats, which were shown by phylogenetic analysis of partial SSU rDNA sequences to cluster with those typically found in primates (Lau et al., 2014). Furthermore there is only one reliable GenBank entry for *E. muris,* isolated from Mongolian gerbil (Kobayashi et al. 2009); the other sequence listed as *E. muris* (FN396613), isolated from *Rattus rattus,* shares 100% identity with *E. coli* ST2, also an 8-nucleated cyst former, and is likely to have been misidentified because of the host. There are no GenBank sequence entries for either *E. funamulae* or *E. citelli*.

Our finding of an *Entamoeba* sequence in DNA extracted from a stool sample from a field vole is significant. The full-length SSU rDNA sequence is here defined as *Entamoeba* RL11 and represents a novel lineage from a host that has not been sampled previously. There is moderate bootstrap support in both NJ (84%) and ML (82%) analyses and strong Bayesian support for placing this new sequence in a clade with *E. muris*. The *Entamoeba* RL11/*E. muris* clade is sister to *Entamoeba* RL7 with high support (> 95%) in all analyses. The clustering of *E. muris* with *Entamoeba* RL7, which has been found in langurs and humans (A. Vidal-Lapiedra pers. commun.), has been previously established (Stensvold et al. 2011). Only future microscopic analysis will confirm if *Entamoeba* RL11 forms 8-nucleated cysts as seen in *E. muris* and *Entamoeba* RL7.

Relatively few rodent hosts have been sampled to date so it is quite likely that more *Entamoeba* diversity remains to be detected. In our experience, rodent stool presents a particular problem unless fresh in that fungi growth on the pellet can potentially mask any *Entamoeba* species sequences present in the extracted DNA, as mentioned earlier for ungulates.

Entamoeba from reptiles

E. invadens produces 4-nucleated cysts and is the most important *Entamoeba* infection of reptiles since it is the causative agent of invasive amebiasis (Geiman and Ratcliffe 1936) and amebic myositis (Chia et al. 2009). Our sequencing here of the complete SSU rDNA of *E. invadens* VK-1 (KR025413), from a Komodo Dragon, identified only a single base difference when compared to the reference strain IP-1, isolated from a snake. In combination with the other existing sequence data (Garcia et al. 2014) and the results from restriction enzyme digestion (Clark and Diamond 1997), this would suggest that intra-specific SSU rDNA sequence diversity in this species is present but low.

A number of other *Entamoeba* species have been identified microscopically in reptiles. These include the 4-nucleated cyst formers *E. insolita* from Galapagos tortoises (Gieman and Wichterman 1937) and *E. terrapinae* from terrapins. An 8-nucleated cyst former, *E. barreti*, has also been reported in Snapping turtle (Geiman and Ratcliffe 1936). In terms of sequence data, the disease-causing agent *E. invadens* constitutes the majority of DNA sequences deposited in GenBank. *E. insolita* (Silberman et al. 1999); *Entamoeba* RL5 from Leopard tortoise (Stensvold et al. 2011) and *Entamoeba* RL6 from Iguana (Silberman et al. 1999) are all represented by single entries. There are no sequence data for *E. barreti*.

In our stool survey, the single python sample proved to be negative for *Entamoeba*, but a novel conditional lineage (*Entamoeba* CL1) was detected in a Giant Aldabran tortoise (Table 1). Unfortunately due to the complexity of the specimen, which also contained *Nyctotherus ovalis* and *E. insolita*, only a partial sequence (470 bp) could be obtained, and it was not possible to assign a new ribosomal lineage number using the published nomenclature criteria (Stensvold et al. 2011). Inclusion of this sequence in our dataset, followed by editing to produce an unambiguous alignment, resulted in a further reduction in the number of useable characters within this sequence (368bp). This in turn resulted in an unresolved and unstable branch that was not specifically linked to any of the other sequences. For this reason, it was decided to exclude this sequence from the final phylogenic analyses.

A study of 127 aquatic turtles (Garcia et al. 2014) found that over half (58.7%) were infected with *Entamoeba*. *E. terrapinae* was found to be the most prevalent species and was present in 63 turtles, followed by *E. invadens* (6 turtles) and *E. moshkovskii* (5 turtles). This represents the first report of *E. moshkovskii* in reptiles. An acknowledged limitation of this study was that it relied on culturing the amoebae rather than direct detection in faecal DNA samples. The same study found a novel putative ribosomal clade represented by three sequences, which we define here as *Entamoeba* CL2.

Our phylogenetic analyses show that the *Entamoeba* CL2 sequence is distinct from *E. terrapinae,* and there is poor bootstrap support in both NJ and ML analyses to support a sister relationship (Fig. 1). The various lineages of reptilian *Entamoeba* appear to form a limited number of clusters in the tree at present (*E. insolita* with *Entamoeba* RL6 and *Entamoeba* RL5, *E. terrapinae* with the newly described lineage) although without strong support. The present as well as prior phylogenetic reconstructions (Stensvold et al. 2011, Stensvold et al. 2010) consistently recover *E. invadens* in a sister relationship with *E. ranarum* (Silberman et al. 1999), the only amphibian-derived *Entamoeba* sequence available to date. As yet, no sequences from reptiles cluster with the 8- or 1-nucleated cyst-producing lineages from mammals, although cysts with eight nuclei have been reported from reptiles on several occasions.

Concluding remarks

Our current survey and phylogenetic reconstructions have further expanded our knowledge of the diversity and host range of *Entamoeba* species. We have discovered new hosts for known *Entamoeba* lineages, the most striking of which is perhaps the reporting of *E. moshkovskii* in cattle and elephants. Furthermore, the finding of four new ribosomal lineages and a further seven ribosomal clades during the course of this work show the merit of the continued sampling of livestock and wild animals.

A notable observation from the present work is the consistently higher incidence of *Entamoeba* infection in animals from managed herds and animals in captivity. In contrast, there was low or no *Entamoeba* positivity in wild animals, even in wild ungulates. It is not clear at present whether the lower rate of *Entamoeba* infection detected in wild animals is an artifact, is due to this group being relatively under-sampled or is real, and only further sampling from wild animal populations can answer this question.

ACKNOWLEDGMENTS

The work described in this manuscript was derived in part from the MSc thesis projects of EJB, ADL and NK. We are very grateful to Dr. Laura Wegener Parfrey for providing us with the *Entamoeba* sequences extracted from her dataset, and for encouraging us to submit the sequences as supplementary information with this manuscript to make them more easily accessible.

LITERATURE CITED

Alfellani, M. A., Taner-Mulla, D., Jacob, A. S., Imeede, C. A., Yoshikawa, H., Stensvold, C. R. & Clark, C. G. 2013. Genetic diversity of *Blastocystis* in livestock and zoo animals. *Protist,* **164:**497-509.

Baines, L., Morgan, E. R., Ofthile, M. & Evans, K. 2015. Occurrence and seasonality of internal parasite infection in elephants, *Loxodonta africana*, in the Okavango Delta, Botswana. *Int J Parasitol: Parasites Wildl,* **4:**43–48.

Bauchop, T. & Martucci, R. W. 1968. Ruminant-like digestion of the langur monkey. *Science (New York, N.Y.),* **161:**698-700.

Chia, M. Y., Jeng, C. R., Hsiao, S. H., Lee, A. H., Chen, C. Y. & Pang, V. F. 2009. *Entamoeba invadens* myositis in a common water monitor lizard (*Varanus salvator*). *Vet Pathol,* **46:**673-676.

Clark, C. G. & Diamond, L. S. 2002. Methods for cultivation of luminal parasitic protists of clinical importance. *Clin Microbiol Rev,* **15:**329-341.

Clark, C. G. & Diamond, L. S. 1997. Intraspecific variation and phylogenetic relationships in the genus *Entamoeba* as revealed by riboprinting. *J Eukaryot Microbiol,* **44:**142-154.

Clark, C. G., Kaffashian, F., Tawari, B., Windsor, J. J., Twigg-Flesner, A., Davies-Morel, M. C., Blessmann, J., Ebert, F., Peschel, B., Le Van, A., Jackson, C. J., Macfarlane, L. & Tannich, E. 2006. New insights into the phylogeny of *Entamoeba* species provided by analysis of four new small-subunit rRNA genes. *Int J Syst Evol Microbiol,* **56:**2235-2239.

Davis, S. D. 1969. Hibernation: intestinal protozoa populations in ground squirrels. *Exp Parasitol,* **26:**156-165.

Diakou, A., Kapantaidakis, E. & Youlatos, D. 2015. Endoparasites of the European ground squirrel (*Spermophilus citellus*) (Rodentia: Sciuridae) in central Macedonia, Greece. *J Nat Hist,* **49:**359-370.

El-Ridi, A. M. S., Ismail, M. M., El-Gamal, R. L. R. & Abdel-Fattah, M. E. M. 1987. Protozoal infections in rodents in Sharkiya Governorate, Egypt. *J Egypt Soc Parasitol,* **17:**275-280.

Fantham, H. B. 1921. Some parasitic protozoa found in South Africa: IV. *South Afr J Sci,* **17:**164–170.

Garcia, G., Ramos, F., Perez, R. G., Yanez, J., Estrada, M. S., Mendoza, L. H., Martinez-Hernandez, F. & Gaytan, P. 2014. Molecular epidemiology and genetic diversity of *Entamoeba* species in a chelonian collection. *J Med Microbiol,* **63:**271-283.

Geiman, Q. M. & Ratcliffe, H. L. 1936. Morphology and life-cycle of an amoeba producing amoebiasis in reptiles. *Parasitology,* **28:**208-228.

Geiman, Q. M. & Wichterman, R. 1937. lntestinal protozoa from Galapagos tortoises (with descriptions of three new species). *J Parasitol,* **23:**331-347.

Heredia, R. D., Fonseca, J. A. & López, M. C. 2012. *Entamoeba moshkovskii:* perspectives of a new agent to be considered in the diagnosis of amebiasis. *Acta Tropica,* **123:**139-145.

Hing, S., Othman, N., Nathan, S. K. S. S., Fox, M., Fisher, M. & Goossens, B. 2013. First parasitological survey of endangered Bornean elephants *Elephas maximus borneensis*. *Endangered Species Research,* **21:**223-230.

Huelsenbeck, J. P. & Ronquist, F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics,* **17:**754–755.

Kinsella, J. M., Deem, S. L., Blake, S. & Freeman, A. 2004. Endoparasites of African forest elephants (*Loxodonta africana cyclotis*) from the Republic of Congo and Central African Republic. *Comp Parasitol,* **71:**104-110.

Kobayashi, S., Suzuki, J. & Takeuchi, T. 2009. Establishment of a continuous culture system for *Entamoeba muris* and analysis of the small subunit rRNA gene. *Parasite,* **16:**135-139.

Lau, Y.L., Jamaiah, I., Rohela, M., Fong, M.Y., Siti, C.O. & Siti, F.A. 2014. Molecular detection of *Entamoeba histolytica* and *Entamoeba dispar* infection among wild rats in Kuala Lumpur, Malaysia. *Trop Biomed*, **31:**721-727.

Leclerc, A., Chavatte, J. M., Landau, I., Snounou, G. & Petit, T. 2014. Morphologic and molecular study of hemoparasites in wild corvids and evidence of sequence identity with *Plasmodium* DNA detected in captive Black-Footed Penguins (*Spheniscus demersus*). *J Zoo Wildl Med,* **45:**577-588.

Matsubayashi, M., Suzuta, F., Terayama, Y., Shimojo, K., Yui, T., Haritani, M., & Shibahara, T. 2014. Ultrastructural characteristics and molecular identification of *Entamoeba suis* isolated from pigs with hemorrhagic colitis: implications for pathogenicity. *Parasitol Res*, **113:**3023-3028.

Mishra, G. S. & Gonzalez, J. P. 1975. Account of a study on the endo parasites if the rat, *Rattus norvegicus* Berkenhout, 1769 , in Tunis *Arch l'Inst Pasteur Tunis,* **52:**71-88.

Nateghpour, M., Motevalli-Haghi, A., Akbarzadeh, K., Akhavan, A. A., Mohebali, M., Mobedi, I. & Farivar, L. 2015. Endoparasites of wild rodents in southeastern Iran. *J Arthropod-Borne Dis,* **9:**1-6.

Neal, R. A. 1948. *Entamoeba histolytica* in wild rats caught in London. *J Hyg (Lond),* **46:**90-93.

Neal, R. A. 1947. *Entamoeba* sp. from the Syrian hamster (*Cricetus auratus*). *Nature,* **159:**502.

Noble, G.A. & Noble, E.R. 1952. Entamoebae in farm mammals. *J Parasitol*, **38:**571-595.

Parfrey, L. W., Walters, W. A., Lauber, C. L., Clemente, J. C., Berg-Lyons, D., Teiling, C., Kodira, C., Mohiuddin, M., Brunelle, J., Driscoll, M., Fierer, N., Gilbert, J. A. & Knight, R. 2014. Communities of microbial eukaryotes in the mammalian gut within the context of environmental eukaryotic diversity. *Frontiers in Microbiology,* **5:**298.

Ray, H. N. & Banik, D. C. 1964. *Entamoeba funambulae* n.sp., from Indian Palm Squirrel, *Funambulus palmarum*. *Bull Calcutta Sch Trop Med,* **12:**114-115.

Royer, T. L., Gilchrist, C., Kabir, M., Arju, T., Ralston, K. S., Haque, R., Clark, C. G. & Petri, W. A., Jr. 2012. *Entamoeba bangladeshi* nov sp., Bangladesh. *Emerg Infect Dis,* **18:**1543-1545.

Shafiyyah, C. O. S., Jamaiah, I., Rohela, M., Lau, Y. L. & Aminah, F. S. 2012. Prevalence of intestinal and blood parasites among wild rats in Kuala Lumpur, Malaysia. *Trop Biomed*, **29:**544-550.

Silberman, J. D., Clark, C. G., Diamond, L. S. & Sogin, M. L. 1999. Phylogeny of the genera *Entamoeba* and *Endolimax* as deduced from small-subunit ribosomal RNA sequences. *Mol Biol Evol,* **16:**1740-1751.

Stensvold, C. R., Lebbad, M. & Clark, C. G. 2010. Genetic characterisation of uninucleated cyst-producing *Entamoeba* spp. from ruminants. *Int J Parasitol,* **40:**775-778.

Stensvold, C. R., Lebbad, M., Victory, E. L., Verweij, J. J., Tannich, E., Alfellani, M., Legarraga, P. & Clark, C. G. 2011. Increased sampling reveals novel lineages of *Entamoeba*: consequences of genetic diversity and host specificity for taxonomy and molecular detection. *Protist,* **162:**525-541.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol*, **28:**2731–2739.

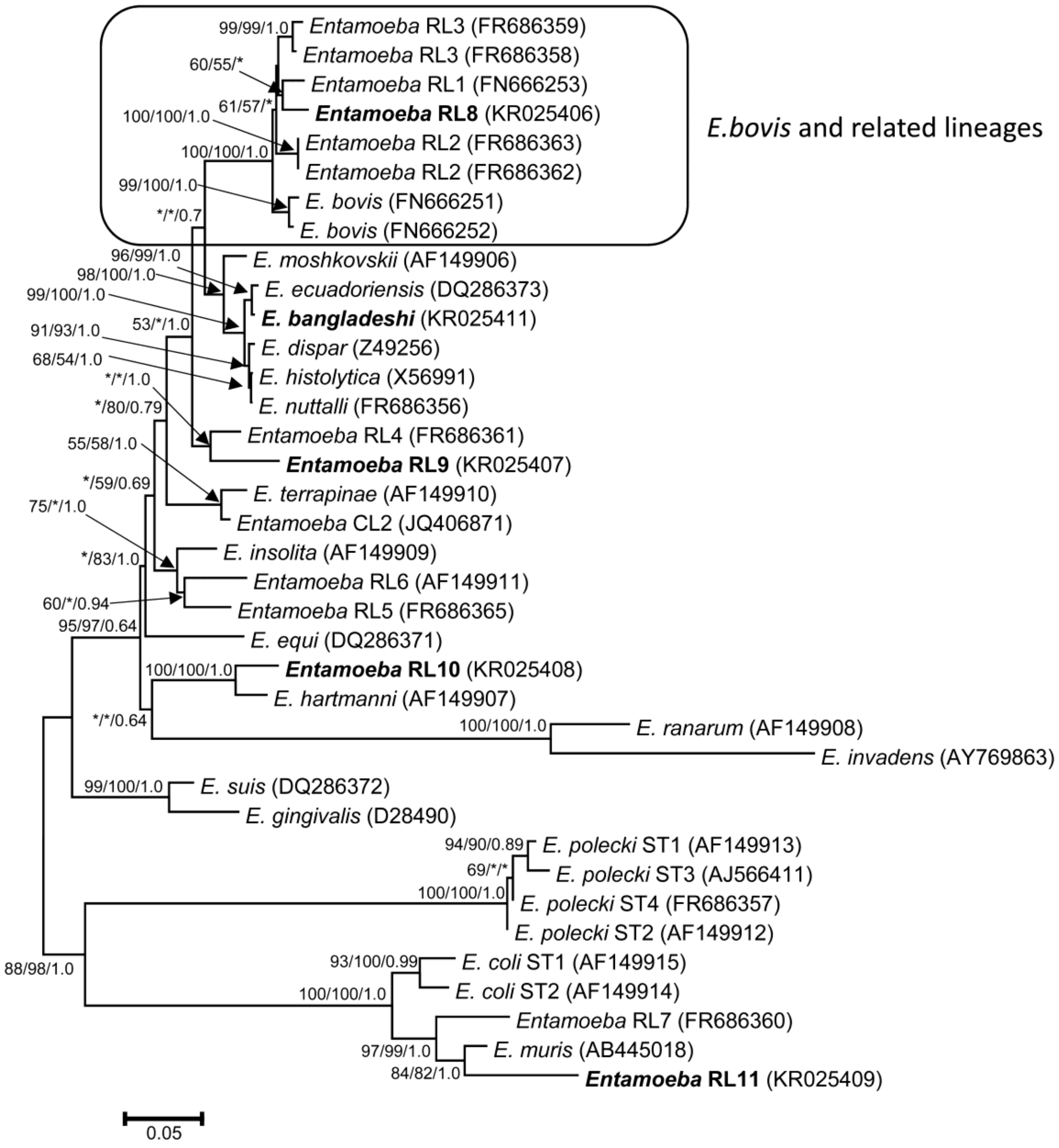
Verweij, J. J. & Stensvold, C. R. 2014. Molecular testing for clinical diagnosis and epidemiological investigations of intestinal parasitic infections. *Clin Microbiol Rev,* **27:**371-418.

Won, Y. S., Jeong, E. S., Park, H. J., Lee, C. H., Nam, K. H., Kim, H. C., Hyun, B. H., Lee, S. K. & Choi, Y. K. 2006. Microbiological contamination of laboratory mice and rats in Korea from 1999 to 2003. *Exp Anim,* **55:**11-16.

FIGURE LEGENDS

Figure 1. The phylogenetic relationships of *Entamoeba* species as inferred from SSU rDNA sequences. New sequences are indicated in bold text. The unrooted distance-based (Neighbor-Joining) tree is shown. Bootstrap proportions and Bayesian posterior probabilities are shown at each node in the order: Neighbor-Joining/Maximum Likelihood/Bayesian analysis. An asterisk indicates a value of less than 50% and if three analyses gave a value of lower than 50% no values are shown for that node. Accession numbers are listed in parentheses. The scale bar represents 0.05 substitutions per site.

Figure 2. A cladogram depicting the phylogenetic relationships among partial *Entamoeba* SSU rDNA sequences from the study of Parfrey et al. (2014). Also included in the cladogram are RL8 from the present study (Cow5) and relevant reference sequences available from GenBank (Cow349, Cow 349.2, Cow350, Cow351, Sheep297, Reindeer100, RoeDeer352, Hulman). The bootstrap consensus maximum likelihood tree is shown as a cladogram for clarity. The corresponding tree showing branch lengths is available as Supplementary Figure S1. Bootstrap support and Bayesian posterior probabilities are shown at each node in the order: Maximum Likelihood/Neighbor-Joining/Bayesian analysis. Only nodes corresponding to existing known species, ribosomal lineages (RL) or newly proposed conditional ribosomal lineages (CL) are labelled. Accession numbers or unique identifier codes are listed in parentheses.


E.bovis and related lineages
99/99/1.0
Entamoeba RL3 (FR686359)
Entamoeba RL3 (FR686358)
60/55/*
Entamoeba RL1 (FN666253)
61/57/*
Entamoeba RL8 (KR025406)
100/100/1.0
Entamoeba RL2 (FR686363)
100/100/1.0
Entamoeba RL2 (FR686362)
99/100/1.0
E. bovis (FN666251)
//0.7
E. bovis (FN666252)
96/99/1.0
E. moshkovskii (AF149906)
98/100/1.0
E. ecuadoriensis (DQ286373)
99/100/1.0
E. bangladeshi (KR025411)
53/*/1.0
91/93/1.0
E. dispar (Z49256)
68/54/1.0
E. histolytica (X56991)
//1.0
E. nuttalli (FR686356)
*/80/0.79
Entamoeba RL4 (FR686361)
55/58/1.0
Entamoeba RL9 (KR025407)
*/59/0.69
E. terrapinae (AF149910)
75/*/1.0
Entamoeba CL2 (JQ406871)
*/83/1.0
E. insolita (AF149909)
Entamoeba RL6 (AF149911)
60/*/0.94
Entamoeba RL5 (FR686365)
95/97/0.64
E. equi (DQ286371)
100/100/1.0
Entamoeba RL10 (KR025408)
E. hartmanni (AF149907)
//0.64
100/100/1.0
E. ranarum (AF149908)
E. invadens (AY769863)
99/100/1.0
E. suis (DQ286372)
E. gingivalis (D28490)
94/90/0.89
E. polecki ST1 (AF149913)
69/*/*
E. polecki ST3 (AJ566411)
100/100/1.0
E. polecki ST4 (FR686357)
E. polecki ST2 (AF149912)
93/100/0.99
E. coli ST1 (AF149915)
88/98/1.0
E. coli ST2 (AF149914)
100/100/1.0
Entamoeba RL7 (FR686360)
97/99/1.0
E. muris (AB445018)
84/82/1.0
Entamoeba RL11 (KR025409)
0.05

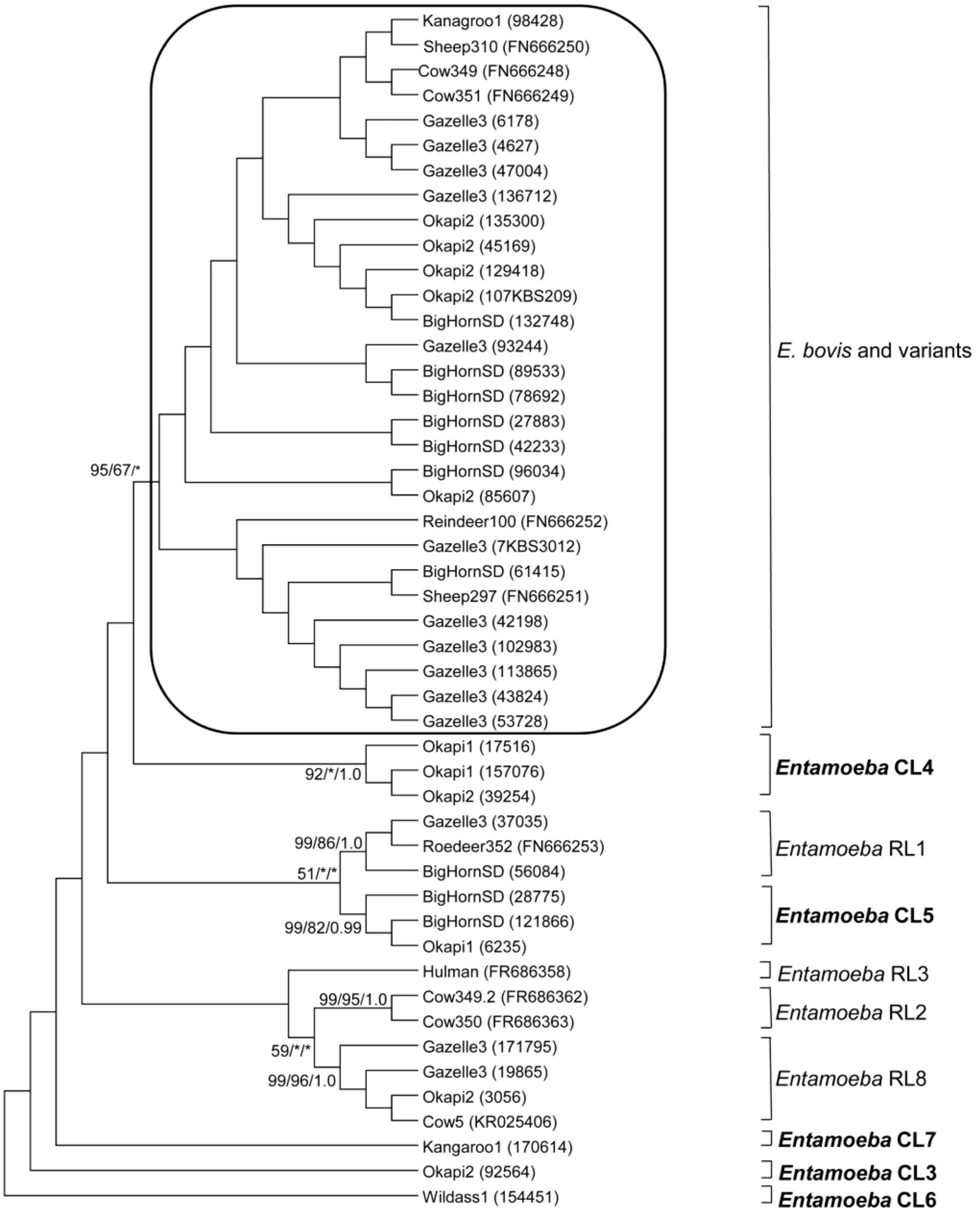

Kanagroo1 (98428)
Sheep310 (FN666250)
Cow349 (FN666248)
Cow351 (FN666249)
Gazelle3 (6178)
Gazelle3 (4627)
Gazelle3 (47004)
Gazelle3 (136712)
Okapi2 (135300)
Okapi2 (45169)
Okapi2 (129418)
Okapi2 (107KBS209)
BigHornSD (132748)
Gazelle3 (93244)
BigHornSD (89533)
BigHornSD (78692)
BigHornSD (27883)
BigHornSD (42233)
95/67/*
BigHornSD (96034)
Okapi2 (85607)
Reindeer100 (FN666252)
Gazelle3 (7KBS3012)
BigHornSD (61415)
Sheep297 (FN666251)
Gazelle3 (42198)
Gazelle3 (102983)
Gazelle3 (113865)
Gazelle3 (43824)
Gazelle3 (53728)
E. bovis and variants
Okapi1 (17516)
92/*/1.0
Okapi1 (157076)
Okapi2 (39254)
Entamoeba CL4
Gazelle3 (37035)
99/86/1.0
Roedeer352 (FN666253)
51/*/*
BigHornSD (56084)
Entamoeba RL1
BigHornSD (28775)
99/82/0.99
BigHornSD (121866)
Okapi1 (6235)
***Entamoeba* CL5**
Hulman (FR686358)
Entamoeba RL3
99/95/1.0
Cow349.2 (FR686362)
Cow350 (FR686363)
Entamoeba RL2
59/*/*
Gazelle3 (171795)
99/96/1.0
Gazelle3 (19865)
Okapi2 (3056)
Cow5 (KR025406)
Entamoeba RL8
Kangaroo1 (170614)
***Entamoeba* CL7**
Okapi2 (92564)
***Entamoeba* CL3**
Wildass1 (154451)
***Entamoeba* CL6**

Table 1. Animal stool samples analyzed during the present study.

Host (species name)	Location	No. of samples	*Entamoeba*-positive samples	*Entamoeba* identified (number of sequences)[a]
Cattle (*Bos taurus*)	Devon, UK	15	4	*E. bovis* (3); *E. moshkovskii* (1)
	Hertfordshire, UK	21	12	*E. bovis* (10); ***Entamoeba* RL8**[b] (1); *Entamoeba* RL4 (1)
Sheep (*Ovis aries*)	Devon, UK	36	13	*E. bovis* (9); *Entamoeba* RL2 (1) ; *Entamoeba* Mixed (3)
	Hertfordshire, UK	19	14	*E. bovis* (14)
Pig (*Sus scrofa domesticus*)	Devon, UK	9	0	
	Vietnam	12	11	*E. polecki* ST1 (7); *E. polecki* mixed ST1 & ST3 (4)
Horse (*Equus ferus caballus*)	Devon, UK	15	3	***Entamoeba* RL9**[b] (3)
Donkey (*Equus africanus asinus*)	Devon, UK	2	0	
Roe deer (*Capreolus capreolus*)	Devon, UK	3	0	
Red deer (*Cervus elaphus*)	Devon, UK	4	0	
Fallow deer (*Dama dama*)	Mauritius	2	2	*E. bovis* (2)
Bank vole (*Myodes glareolus*)	Devon, UK	7	0	
Field vole (*Microtus agrestis*)	Northumberland, UK	12	1	***Entamoeba* RL11**[b] (1)

Wood mouse (*Apodemus sylvaticus*)	Devon, UK	5	0	
Stoat (*Mustela erminea*)	Devon, UK	2	0	
European Badger (*Meles meles*)	Devon, UK	1	0	
Chicken (*Gallus gallus domesticus*)	Devon, UK	1	0	
Fox (*Vulpes vulpes*)	Devon, UK	1	0	
Rabbit (*Oryctolagus cuniculus*)	Devon, UK	1	0	
Asian Elephant (*Elephas maximus*)	Amsterdam, The Netherlands	4	1	***Entamoeba* RL10**[b] (1)
Goose (*Anser domesticus*)	Devon, UK	2	0	
Royal python (*Python regius*)	London, UK	1	0	
Aldabran Giant tortoise (*Aldabrachelys gigantea*)	Mauritius	3	1[c]	***Entamoeba* CL1**[b] (1); *E. insolita* (1)
Total		178	62	

Numbers in parentheses indicate the number of samples corresponding to particular species/ribosomal lineages.

[a] Bold typeface indicates either a new ribosomal lineage (RL) or conditional lineage (CL) identified during in this study.

[b] Accession numbers for newly described RLs and CL are as follows: *Entamoeba* RL8 (KR025406), *Entamoeba* RL9 (KR025407), *Entamoeba* RL10 (KR025408), *Entamoeba* RL11 (KR025409) and *Entamoeba* CL1 (KR025410).

[c] Two distinct *Entamoeba*s were present in this one sample.

Table 2. List of animals surveyed by Parfrey et al. (2014) in which *Entamoeba* sequences were found

Host (species name)	Location	Number of samples	*Entamoeba*-positive samples[a]	*Entamoeba* identified (number of sequences)[b]
Wild ass (*Equus asinus*)	Saint Louis zoo, USA	1	1	***Entamoeba* CL6** (1)
Zebra (*Equus grevyi*)	Saint Louis zoo, USA	2	1	*E. hartmanni* (1)
Gazelle (*Gazella spekei*)	Saint Louis zoo, USA	2	2	*E. bovis* (11); *Entamoeba* RL1 (1); *Entamoeba* RL8 (2)
African bush elephant (*Loxodonta africana*)	Namibia	1	1	*E. moshkovskii*
Red kangaroo (*Macropus rufus*)	Saint Louis zoo, USA	2	1	*E. bovis* (1); ***Entamoeba* CL7** (1)
Okapi (*Okapia johnstoni*)	Saint Louis zoo, USA	3	2	*E. bovis* (5); ***Entamoeba* CL3** (1); ***Entamoeba* CL4** (3); ***Entamoeba* CL5** (1); *Entamoeba* RL8 (1);
Bighorn sheep (*Ovis canadensis*)	Saint Louis zoo, USA	2	1	*E. bovis* (7); *Entamoeba* RL1 (1); ***Entamoeba* CL5** (2)
Baboon (*Papio hamadryas*)	Namibia	1	1	*E. hartmanni*
Sumatran oranutan (*Pongo abelii*)	Saint Louis zoo, USA	1	1	*E. hartmanni*
Total		**15**	**11**	

[a] Several samples contained more than one *Entamoeba* species or lineage

[b] Bold typeface in this column indicates a conditional lineage (CL). These are defined as distinct sequences that are too short to meet the criteria as established for ribosomal lineages (Stensvold et al. 2011)

SUPPORTING INFORMATION

Supplementary Data File S1. Partial *Entamoeba* SSU rDNA sequences from the study of Parfrey et al. (2014). *Entamoeba* sequences extracted from the curated SILVA 108 database (http://qiime.org/home_static/dataFiles.html) generated in Parfrey et al. (2014) and used to obtain Fig. 2 are listed.

Supplementary Figure S1. A phylogram depicting the phylogenetic relationships among partial *Entamoeba* SSU rDNA sequences from the study of Parfrey et al. (2014). This is the same tree as shown in fig. 2, except with branch lengths shown.

Supplementary Data File S2. Distance matrix. The estimated distances between sequences in the Neighbor-Joining tree shown in fig. 1 were computed using the Maximum Composite Likelihood method.

Supplementary data file S1.

Partial *Entamoeba* SSU rRNA gene sequences from the study of Parfrey et al. (2014). *Entamoeba* sequences extracted from the curated SILVA 108 database (http://qiime.org/home_static/dataFiles.html) generated in Parfrey et al. (2014) and used to obtain Fig. 2 are listed.

>Okapi2 (45169)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
TAAAGCGCTTAGCTTGCGGGTGCCCTGCTCTGCGGGGAGGAAGCTTGCGATAAACGGCGC
GGAGGCGATGCCGGTTTCGGCCGGTGTCATTACTTTGAAAAAATAGGGTGTTCAAAGCA
AATCTTATGTTAATGAATAATGAAGCATGGGACAATATCGAGGAGATCTTTCGGGATTT
CGGGAAAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATGACGGGAGAGGTAAAA
TTCCATGATCGCCATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTAA
TCAAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACTATAA
ACGATGTCAACCAAGGATTGGATTAGTTTTAGAGTGACAGAAGTCCGGTAACGCTGTTA
CTGGGTTGACGGATCTCGCTTCCACCTTATTCAGAACTTAAAGAGAAATCTTGAGTT

>Gazelle3 (42198)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
TAAAGCGCTTAGCGTATGGGTGCCCTGCTCGTCGGGGAGGAAATGCGCATCAAACGGTGC
GAAGGCGATGCCGGTTTCGATCGGTGTCATTACTTTGAAAAAATAGGGTGTTCAAAGCA
AATCTTATGTTAATGCATAATGAAGCATGGGACAATATCGAGGAGATCCTTCGGGATTT
CGGGATAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATGACGGGAGAGGTAAAA
TTCCATGATCGCCATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTAA
TCAAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACTATAA
ACGATGTCAACCAAGGATTGGATTAGTTTAAGAGGGACAGAATTCCGGCAACGCTGTTG
TTGGGATGACGGACCTCGCTTCCACCTTATTCAGAACTT

>Gazelle3 (37035)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
TAATGTTTCCTAGTATATGGAGGTTCTGCTTGCAGGGCTGATATGTACAATTAATGGAA
CTTAGGCGATGCTGCTCACGCGGTGTCATTACTTTGAAAAAATAGGGTGTTTAAAGCAA
ATCTTATGTTAATGAATAATGAAGCATGGGATAATATTGAGGAGATTCTTCGGGATTTC
GAGAGAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATAACGGGAGAGGTAAAA
TTCCATGATCGCTATAGGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTAA
TCAAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACTATAA
ACGATGTCAACCAAGGATTGGATTAGTATTTGGACGACAGAGGTTAGATAACATTGTTA
TTTGACTAACGGAGT

>Wildass1 (154451)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
AAACGTGTATTTGTTTTAGGAAAGTTCTCAGAACTGGAATAAAACGCTTAAAGTACACC
GAAGGAGATGAAATGAGCAATCATTTTATCATTACTTTGAAAAAATAGAGTGTTTCAAG
CAAAACATTTTGTTAATGAATAATGAAGCATGGGACAATGCTGAGGAGACGGTCTTCGG
ACTGTTTCGAGATAAGGATTAAAGGAATACTTGGG

>BigHornSD (96034)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
TAAAGCGCTTAGCATATGGGTGCCCTGCTCTGTGGGGAGGAAGTATGCGCTGAACGGTGC
GAAGGCGACGTCGGTTTCGGCCGATGTCATTACTTTGAAAAAATAGGGTGTTCAAAGCA
AATCTTATGTTAATGCATAATGAAGCATGGGACAATATCGAGGAGATCCTTCGGGATTT

CGGGAAAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAACGACGGGAGAGGTAAAA
TTCCATGATCGCCATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTAA
TCAAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACTATAA
ACGATGTCAACCAAGGATTGGATTAGTTTTAGGGTGACAGAAGTCCGGTAACGCTGTTA
CTGGGTTGACGGATCTCGCTTCC
>Okapi2 (135300)
TTCCAGCTCCAATAGTGTATATTAAAGTCGCTGTGATTAAAACGCTCGTAGTTGAATTAT
GAAGCGCTTAGCTTGCGGGTGCCCTGCTCTGCGGGGAGGAAGCTTGCGATAAACGGCGCG
GAGGCGATGCCGGTTTCGGCCGGTGTCATTACTTTGAAAAAATAGGGTGTTCAAAGCAA
ATCTTATGTTAATGAATAATGAAGCATGGGGCAATATCGAGGAGATCTTTCGGGATTTC
GGGAAAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATGACGGGAGAGGTAAAAT
TCCATGATCGCCATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTAAT
CAAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACTATAAA
CGATGTCAACCAAGGATTGGATTAGTTTTAGAGTGACAGAAGTCCGGTAACGCTGTTAC
TGGGTTGACGGATCTCGCTTCCACCTTATTCAGACTTAAAGAGAATC
>Gazelle3 (102983)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAAACGCTCGTAGTTGAATT
ATAAAGCGCTTAGCGTATGGGTGCCCTGCTCGGTCGGGGAGGAAATGCGCATCAAACGGT
GCGAAGGCGATGCCGGTTTCGATCGGTGTCATTACTTTAGAAAAAATAGGGTGTTCAAA
GCAAATCTTATGTTAATGCATAATGAAGCATGGGACAATATCGAGGAGATCCTTCGGGA
TTTCGGGATAAGGATTAAGAGGAGCAATTGGGGTGATTCAGAAAATGACGGGAGAGGTA
AAATTCCATGATCGCCATAAGATGCACGAGAGCGAAAGCATTTCGCTCAATTGCGTTCAT
TAATCAAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACTA
TAAACGATGTCAACCAAGGATTGGATTAGTTTTAGAGGAAGTGTGACTACAAATCATTG
ATTTGCTGGAACACCGACTCGCTACTACCTCATTCAGAACTTAATGAGAAA
>BigHornSD (89533)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
TAGAGCGCTTAGCATGCGGGTGCCCTGCTCTGCGGGGAGGAAGTATGCGGTAAACGGCGC
GGAGGCGATGCCGGTTTCGGCCGGTGTCATTACTTTGAAAAAATAGGGTGTTCAAAGCA
AATCTTATGTTAATGCATAATGAAGCATGGGACAATATTGAGGGGTTCTTTCGGGAATT
CGAGATAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATAACGGGAGAGGTAAAA
TTCCATGATCGCTATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTAA
TCAAGAACGAAAGTTAGGGGATCGAAGACGATCAGATACCGTCGTAGTCCTAACTATAA
ACGATGTCATCCAAGGATTGGATTAGTTTTAGAGAGACAGATGCTCGGTAACATCGTTA
TTGAGCGGACGGATCTCGCCTCCACCTTATTCAGAACTTAAAG
>Okapi2 (92564)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTG
AAAAGTGTCTAGCTTGGGAAAGCCGGGTAACCGGGGGACTCTTGCGTTAACGGACATGA
AGGTTATGGTCGCAAGGCCGTAATTACTTTGAAAAAATAGGGTGTTTAAAGCAAATCTT
ATGTTAATGAATAATGAAGCATGGGACAATATTGAGGAGATCCCTTTTGGGATGTCGAG
ATAAGGATTAAAAGGAACAATTGGGGTGATTCAGAAAATAACGGGAGAGGTGAAAATCC
ATGATCGCTATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTAATCA
AGAACGAAAGTTAGGGGATCGAAGACGATCAGATACCGTCGTAGTCCTAACTATAAACG
ATGTCAACCAAGGATTGGATTAGTTTTA
>BigHornSD (56084)

TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
TAATGTTTCCTAGTATATGGAGGTTCTGCTTGCAGGGCTGATATGTACAATTAATGGAA
CTTAGGCGATGCTGCTCACGCGGTGTCATTACTTTGAAAAAATAGGGTGTTTAAAGCAA
ATCTTATGTTAATGAATAATGAAGCATGGGACAATATTGAGGAGATTCTTCGGGATTTC
GAGAGAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATAACGGGAGAGGTAAAA
TTCCATGATCGCCATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTAA
TCAAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACTATAA
ACGATGTCAACCAAGGATTGGATTAGTTTAAGAGGGACAGAAGTCCGGCAACGCTGTTG
TTGGGCTGACGGACCTCGCT
>Gazelle3 (7KBS3012)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
TAAAGCGCTTAGCGTATGGGTGCCCTGCTCGTCGGGGAGGAAATGCGCATCAAACGGTGC
GAAGGCGATGCCGGTTTCGATCGGTGTCATTACTTTGAAAAAATAGGGTGTTCAAAGCA
AATCTTATGTTAATGAATAATGAAGCATGGGATAATATTGAGGAGATTCTTTGGGATTT
TCGAGAGAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATAACGGGAGAGGTAAA
ATTCCATGATCGCTATAGGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTA
ATCAAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACTATA
AACGATGTCAACCAAGGATTGGAT
>BigHornSD (121866)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
TAATGGACCCTGTATCTGGAAGCTCTGCTTGCGGGGTTGAACGATACGTTATACGGTTCT
TAGGCGATGCTCGTTAACGCGAGTGTCATTACTTTGAAAAAATAGGGTGTTTAAAGCAA
ATCTTATGTTAATGAATAATGAAGCATGGGACAATATTGAGGGGTTCTTTCGGGAATTC
GAGATAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATAACGGGAGAGGTAAAAT
TCCATGATCGCTATAAGATGCACGAGAACGAAAGCATTTCACTCAATTGCGTTCATTAAT
CAAGAACGAAAGTTAGGGGATCGAAGACGATCAGATACCGTCGTAGTCCTAACTATAAA
CGATGTCAACCAAGGATTGGATTAGTTTTAGAGAGACAGATGCTCGGTAACATCGTTAT
TGAGCGGACGGATCTCGCCTCCACCTTATTCAGAACTTAAAGAGAAATCTT
>Okapi1 (17516)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
TAAAGCGTTTAGTGCGTGGAGCACTGCTTGCAGTGTGAAGGGCACAATCAAACAACGCCT
TAGGTGATGCTCTTCGGAGTGTCATTACTTTGAAAAAATAGGGTGTTTAAAGCAAATCT
TATGTTAATGAATAATGAAGCATGGGACAATATTGAGGAGATCCTTCGGGATTTCGAGA
AAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATAACGGGAGAGGTAAAATTCCA
TGATCGCTATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTAATCAA
GAACGAAAGTTAGGGGATCGAAGACGATCAGATACCGTCGTAGTCCTAACTATAAACGA
TGTCAACCAAGGATTGGATTAGTTAAAGAGTGACAGATGCCGAATAGTATGGCTATTCT
GGCGGACGGAACTCG
>Gazelle3 (19865)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
TAAGACACTAGAATGGGCAGAAGCTTTTCGAAGTGGAGCTCTATTCGCTAAAGTGTTTA
GGTGATACTTACTTCGGTTAGTGTCATTACTTTGAAAAAATAGGGTGTTTAAAGCAAAT
CTTATGTTAATGAATAATGAAGCATGGGACAATATTGAAGGAGATCCTTTATGGATTTC
GAGGATAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATAACGGGAGAGGTAAA
ATTCCATGATCGCTATACGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTA
ATCAAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACTATA

```
AACGATGTCAACCAAGGATTGGATTAGTTTAGGAGGGACAGAATTCCGGCAACGCTGTT
GTTGGGATGACGGACCTCGCTTCCACCTTATTCAGAACTTAAAGAGAAATCTCGAGTT
>Okapi2 (85607)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
TAAAGCGCTTAGCATATCGGTGCCCTACTCTGTGGGGAGGAAGTATGCGCTGAACGGCGC
GAAGGCGATGTCGGTTTCGGCCGATGTCATTACTTTGAAAAAATAGGGTGTTCAAAGCA
AATCTTATGTTAATGAATAATGAAGCATGGGACAATATCGAGGAGATCCTTCGGGATTT
CGGGATAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATGACGGGAGAGGTAAAA
TTCCATGATCGCCATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTAA
TCAAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACTATAA
ACGATGTCAACCAAGGATTGGATTAGTCATAGAGGGACAGAAGTCCGGCAACGCTGTTG
TCGGGCTGACGGACCTCGCTTCCACCTTATTCAGAACTTAAAGAGAAAT
>Okapi1 (157076)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
TAAAGCATTTAGTTCTTGGAGGCTCTGTTCGCAGGGTTGAAGAGGACAACAAACGAATG
CCTTAGGTGATGCTCTTCGGAGTGCCATTACTTTGAAAAAATAGGGTGTTTAAAGCAAA
TCTTATGTTAATGAATAATGAAGCATGGGACAATATCGAGGAGATCTTTCGGGATTTCG
AGAAAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATAACGGGAGAGGTAAAAT
TCCATGATCGCTATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTAAT
CAAGAACGAAAGTTAGGGGATCGAAGACGATCAGATACCGTCGTAGTCCTAACTATAAA
CGATGTCAACCAAGGATTGGATTAGTTATAGAGTGACAGATGCCGGGTAGCATGGCTAT
CAGGCGGACGGAGCTCGC
>Gazelle3 (43824)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
TAAAGCGCTTAGCGTATGGGTGCCCTGCTTGTCGGGGAGGAAATGCGCATCAAACGGTGC
GAAGGCGATGCCGGTTTCGATCGGTGTCATTACTTTGAAAAAATAGGGTGTTCAAAGCA
AATCTTATGTTAATGCATAATGAAGCATGGGACAATATCGAGGAGACCCTTCGGGTTTC
GGGTAGTAA
>Gazelle3 (113865)
TTCCAGCTCCAATAGTGTATATTAAAGTTAGCTGTGATTAAAACGACTCGTAGTTGAAT
TATAAAGCGCTTAGCGTATGGGTGCCCTGCTTGTCGGGGAGGGAAATGCGCATCAAACGG
TGCGAAGGCGATGCCGGTTTCGATCGGTGTCATTACTTTGAAAAAATAAGGGTGTTCAA
AGCAAATCTTATGTTAATGCATAATGAAGCATGGGACAATATCGAGGAGATCCTTCGGG
ATTTCGGGATAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATGACGGGAGAGGT
AAAATTCCATGATCGCCATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCG
TTAATCAAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACT
ATAAACGATGTCAACCAAGGATTGGATTAGTTTTAA
>Kangaroo1 (170614)
TTCCAGCTCCAATAGTGTATATTAAAGTTGTTGTGATTAAAACGCTCGTAGTTGAATTA
TTAGACAGTCTTATTTAACGGATTTTGAGGTTCGCCGAAAATGAAGTTAATACCAAATG
GACTGTTGAAGGCGATATGTCGTTTACGACATGTCATTACTTTGAATAAATTAAGGTGT
TTAAAGCAAAACTTTTGTTAATGAATAATTAAGCATAGGACAATATTAAGGAGATACTT
AGGTATTTCGAGAGAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATAACGGGAG
AGGTAAAATTCTAAGATCGCTATAAGATGAACGAGAGCGAAAGCATTTCACTCAATTGT
GTCCATTAATCAAGAACGAAAGTTAGGGGATCGAAGACGATCAGATACCGTCGTAGTCC
TAACTATAAACGATGTC
```

>Okapi2 (3056)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
TAAGACACTAGAATGGGCAGAAGCTTTTCGAAGTGGAGCTCTATTCGCTAAAGTGTTTA
GGTGATACTTACTTCGGTTAGTGTCATTACTTTGAAAAAATAGGGTGTTTAAAGCAAAT
CTTATGTTAATGAATAATGAAGCATGGGACAATATTGAAGGAGATCCTTTATGGATTTC
GAGGATAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATAACGGGAGAGGTAAA
ATTCCATGATCGCTATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTA
ATCAAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACTATA
AACGATGTCAACCAAGGATTGGATTAGTTTTTGAGTGACTAAGAGTTAATAGCATTGCT
ATTGACTTGAAGGAACTCGCTTTCACCTTATTCAGAACTTAAAGAGAAATCTT

>BigHornSD (27883)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATGA
TAAAGCGCTTAGCATACGGGTGCTCTGCTCTGCGGGGAGGAAGTATGCAGTAGACGGCGC
GGAGGCGATGCCGGTTCCGGTCGGTGTCATTACTTTGAAAAAATAGTGTGTTCAAAGCA
AATCTTATGTTAATGAATAATGAAGCATGGGACAATATCGAGGAGATCCTTCGGGATTT
CGGGAAAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATGACGGGAGAGGTAAAA
TTCCATGATCGCCATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCATTCATTAA
TCAAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACTATAA
ACGATGTCAACCAAGGATTGGATTAGTTTTAGAGTGACAGAGGTCCGGTAACACAGTTA
CTGGGCTGACGGATCTCGCTTCCACCTTATTCAGAACTTAA
>BigHornSD (42233)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATGA
TAAAGCGCTTAGCATACGGGTGCTCTGCTCTGCGGGGAGGAAGTATGCAGTAGACGGCGC
GGAGGCGATGCCGATTCCGGTCGGTGTCATTACTTTGAAAAAATAGTGTGTTCAAAGCA
AATCTTATGTTAATGAATAATGAAGCATGGGACAATATCGAGGAGATCCTTCGGGATTT
CGGGATAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATAACGGGAGAGGTAAAA
TTCCATGATCGCTATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTAA
TCAAGAACGAAAGTTAGGGGATCGAAGACGATCAGATACCGTCGTAGTCCTAACTATAA
ACGATGTCAACCAAGGATTGGATTAGTTTTAGAGAGACAGATGCTCGGTAACATCGTTA
TTGAGCGGACGGATCTCGCCTCCACCTTATTCAGAACTTAAAGAGAAATCTTGAGTTT
>Gazelle3 (53728)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
TAAAGCGCTTAGCGTATGGGTGCCCTGCTCGTCGGGGAGGAAATGCGCATCAAACGGTGC
GAAGGCGATGCCGGTTTCGATCGGTGTCATTACTTTGAAAAAATAGGGTGTTCAAAGCA
AATCTTATGTTAATGCATAATGAAGCATGGGACAATATCGAGGAGATCCTTCGGGATTT
CGGGATAA
>BigHornSD (78692)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
TAAAGCGCTTAGCATGCGGGTGCCCTGCTCTGCGGGGAGGAAGTATGCGGTAAACGGCGC
GGAGGCGATGCCGGTTTCGGCCGGTGTCATTACTTTGAAAAAATAGGGTGTTCAAAGCA
AATCTTATGTTAATGCATAATGAAGCATGGGACAATATCGAGGAGATCCTTCGGGATTT
CGGGAAAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATGACGGGAGAGGTAAAA
TTCCATGATCGCTATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTAA
TCAAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACTATAA

ACGATGTCAACCAAGGATTGGATTAGTATTTGGACGACAGAGGTTAGATAACATTGTTATTTGACTAACGGAGTCCGCTT

>Okapi2 (107KBS209)

TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTATAAAGCGCTTAGCTTGCGGGTGCCCTGCTCTGCGGGAGGAAGCTTGCGATAAACGGCGCGGAGGCGATGCCGGTTTCGGCCGGTGTCATTACTTTGAAAAAATAAGGGTGTTCAAAGCAAATCTTATGTTAATGAATAATGAAGCATGGGACAATAATCGAGGAGATCTTTCGGGATTTTCGGGAAAAGGAT

>Gazelle3 (6178)

TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATGATAAAGCGCTTAGCATGCGGGTGCGCTGCTTCGCAGTGAGGAAGTATGCGATTAATGGCGCGAAGGCGATGCTGGTTTCGGCCGGTGTCATTACTTTGAAAAAATAGGGTGTTCAAAGCAAATCTTATGTTAATGAATAATGAAGCATGGGACAATATTGAGGAGATCCTTCGGGATTTCGGGAAAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATGACGGGAGAGGTAAAATTCCATGATCGCCATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTAATCAAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACTATAAACGATGTCAACCAAGGATTGGATTAGTTTTAGAGTGACAGAAGTCCGGCAACGCTGTTGTTGGGCTGACGGA

>Okapi1 (6235)

TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGCAGTTGAATTATAATGGACCCTGTATCCGGAGGCCCTGCTTGCGGGGTTAAAGATACGTTATACGGTTCTTAGGCGATGCTCGTTAACGCGAGTGTCATTACTTTGAAAAAATAAGGGTGTTTAAAGCAAATCTTATGTTAATGAATAATGAAGCATGGGACAATATTGAGGAGTTCCTTCGGGAATTCGAGATAAGGATTAAGAGGAACAATTGGGTGATTCAGAAAATAACGGGAGAGGTAAAACTCCATGATCGCTATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTAATCAAGAACGAAAGTTAGGGGATCGAAGACGATCGGATACCGTCGTAGTCCTAACTATAAACGATGTCAACCAAGGATTGGATTAGTTTTAGAGAGACAGACGCTCGGTAGCATTGTTATTGAGCTGACGGATCTCGCCTCCACCTTATTCAGAACTTAAAGAGAAATCTTGAAGTTTATGGA

>Gazelle3 (4627)

TTCCAGCTCCAATAGTGTATATTAAAGTTAGCTGTGATTAAAACGACTCGTAGTTGAATGATAAAGACGCTTAGCATGCGAGTGCGCTGCTTCGTCAGTGAGGAAGTATGCGATTAATGGCGCGAAGGCGATGCTGGTTTCGGTCCGGTGTCATTACTTTGAAAAAATAA

>Gazelle3 (171795)

TTCCAGCTCCAATAGTGTATATTAAAGTTAGCTGTGATTAAAACGACTCGTAGTTGAATTATAAGACACTAGAATGGGCAGAAGCTTTTCGTAAGTGGAGCTCTATTCGCTAAAGTGTTTAGGTGATACTTACTTCGGTTAGTGTCATTACTTTGAAAAAATAAGGGTGTTTAAAGCAAATCTTATGTTAATGAATAATGATGCATGGGACAATATTGAAGGAGATCCTTTATGGATTTCGAGGATAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAATAACGGGAGAGGTAAAATTCCATGATCGCTATAAGATGCACGAGAGCGAAAGCATTTCACTCGATTGCGTTCATTAATCAAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACTATAAACGATGTCAACCAAGGATTGGATTAGTATTTGGACGACAGAGGTTAGATAACATTGTTATT

>Okapi2 (39254)

TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
TAAAGCATTTAGTGCTTGGAGGCTCTGTTCGCAGGGTTGAAGAGTACAACAAACGAATG
CCTTAGGTGATGCTCTTCGGAGTGTCATTACTTTGAAAAAATAGAGTGTTTAAAGCAAA
TCTTATGTTAATGAATAATGAAGCATGGGACAATATCGAGGAGATCTTTCGGGATTTCG
AGAAAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATAACGGGAGAGGTAAAAT
TCCATGATCGCCATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTAAT
CAAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACTATAAA
CGATGTCAACCAAGGATTGGATTAGTTTTAGAGTGACAGAAGTCCGGTAACGCTGTTAC
TGGGTTGACGGATCTCGCTTCCACCTTATTCAGAACTTAAAGAGAAATCTTGAG

>BigHornSD (28775)

TTCCAGCTCCAATAGTGCATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTAT
AATGGACCCTGTATCTGGAAGCTCTGCTTGCGGGGTTGAACGATACGTTATACGGTTCTT
AGGCGATGCTCGTTAACGCGAGTGTCATTACTTTGAAAAAATAGGGTGTTTAAAGCAAA
TCTTATGTTAATGAATAATGAAGCATGGGACAATATTGAGGAGATTCTTCGGGATTTCG
AGAGAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATAACGGGAGAGGTAAAATT
CCATGATCGCTATAGGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTAATC
AAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACTATAAAC
GATGTCAACCAAGGATTGGATTAGTATTTGGACGACA

>Kangaroo1 (98428)

TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
TAAAGCGCTTAGTATGCGGGAGCTCCGCTTTGCGGGGAAGAAGTATGCGATAAACGGCG
CGGAGGCGATGCCGGTTTCGGTCGGTGTCATTACTTTGAAAAAATAGGGTGTTCAAAGC
AAATCTTATGTTAATGAACAATGAAGCATGGGACAATATCGAGGAGATCCTTCGGGATT
TCGGGAAAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATGACGGGAGAGGTAAA
ATTCCATGATCGCCATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTA
ATCAAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACTATA
AACGATGTCAACCAAGGATTGGATAAATTTTAAAAACTGAGAGTGATTAAAGCATTGCT
TTAGGATCTTAAAGTTTTATTACTACC

>Okapi2 (129418)

TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
TAAAGCGCCTAGCTTGCGGGTGCCCTGCTCTGCGGGGAGGAAGCTTGCGATAAACGGCGC
GGAGGCGATGCCGGTTTCGGCCGGTGTCATTACTTTGAAAAAATAGGGTGTTCAAAGCA
AATCTTATGTTAATGAATAATGAAGCATGGGACAATATCGAGGAGATCTTTCGGGATTT
CGGGAAAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATGACGGGAGAGGTAAAA
TTCCATGATCGCCATAAGATGCACGAGGGCGAAAGCATTTCACTCAATTGCGTTCATTAA
TCAAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACTATAA
ACGATGTCAACCAAGGATTGGATTAGTTTTAGAGGAAGTGTGACTACAAATCATTGATT
TGCTGGAACACCGACTCGCTACTACCTCATTCAGAACTTAATGAGAAATCA

>Gazelle3 (47004)

TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATGA
TAAAGCGCTTAGCATGCGGGTGCGCTGCTTCGCAGTGAGGAAGTATGCGATTAATGGCGC
GAAGGCGATGCTGGTTTCGGCCGGTGTCATTACTTTGAAAAAATAGGGTGTTCAAAGCA
AATCTTATGTTAATGAATAATGAAGCATGGGACAATATTGAGGAGATCCTTCGGGATTT
CGGGAAAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATGACGGGAGAGGTAAAA
TTCCATGATCGCCATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTAA
TCAAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCCAACTATAA

ACGATGTCAACCAAGGATTGGATTAGTTTTTGAGTGACTAAGAGTTAATAGCATTGTTA
TTTGACTAACGGAG

>Gazelle3 (93244)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGACTCGTAGTTGAATT
ATAAAGCGCTTAGCATGCGGGTGCCCTACTCTGCGGGGAGGAAGTATGCGGTAAACGGCG
CGGAGGCGATGCCGGTTTCGGCCGGTGTCATTACTTTGAAAAAATAAGGGTGTTCAAAG
CAAATCTTATGTTAATGAATAATGAAGCATGGGACAATATTGAGGAGATCCTTCGGGAT
TTCGGGAAAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATGACGGGAGAGGTAA
AATTCCATGATCGCCATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATT
AATCAAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACTAT
AAACGATGTCAACCAAGGATTGGATTAGTTTTAGAGTGACAGAAGTCCGGCAACGCTG
TTGTTGGGCTGACGGATCTCGCTTCCACCTTATTCAGAACTTAAAGAGAAATCTCGAGTT
TATG

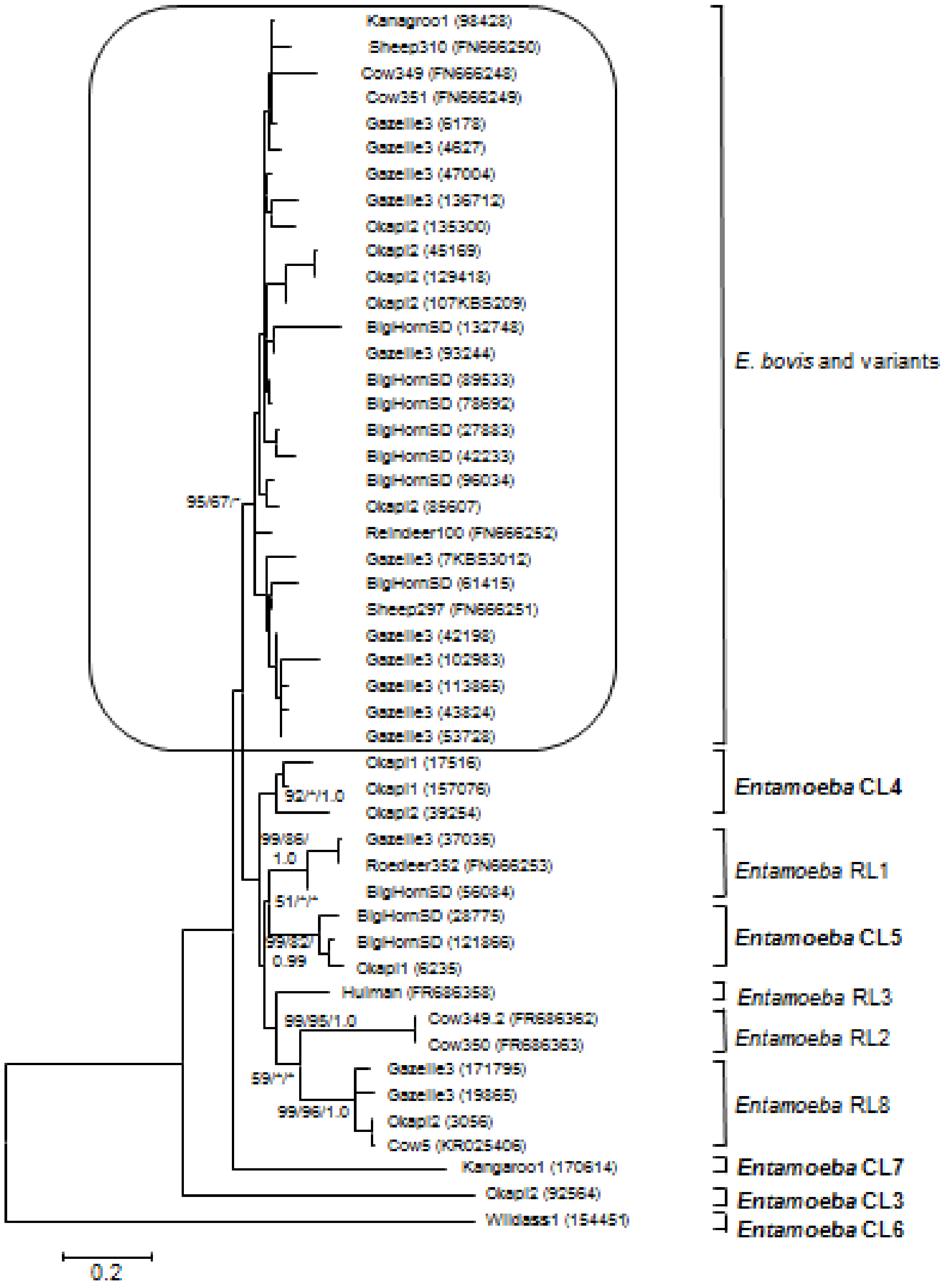
>BigHornSD (61415)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
TAAAGCGCTTAGCGTATGGGTGCCCTGCTCGTCGGGGAGGAAATACGCATCAAACGGTGC
GAAGGCGATGCCGGTTTCCGCCGGTGTCATTACTTTGAAAAAATAGGGTGTTCAAAGCA
AATCTTATGTTAATGCATAGTGAAGCATGGGACAATATCGAGGAGATCCTTCGGGATTT
CGGGATAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATGACGGGAGAGGTAAAA
TTCCATGATCGCCATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTAA
TCAAGAACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACTATAA
ACGATGTCAACCAAGGATTGGATTAGTATTTGGACGACAGAGGTTAGATAACAATTGTT
ATTTGACTAA

>Gazelle3 (136712)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
TAAAGCGCTTAGCTTGCGGGTGCCCTGCTCTGCGGGGAGGAAGCTTGCGATAAACGGCGC
GGAGGCGATGCCGGTTTCGGCCGGTGTCATTACTTTGAAAAAATAGGGTGTTCAAAGCA
AATCTTATGTTAATGAATAATGAAGCATGGGACAATATCGAGGAGATCCTTCGGGATTT
CGGGATAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATGACGGGAGAGGTAAAA
TTCCATGATCGCCATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTAA
TCAAGGACGAAAGTTAGGGGATCGAAGACGATTAGATACCGTCGTAGTCCTAACTATAA
ACGATGTCAACCAAGGATTGGATTAGTTTAAGAGGGACAGAATTCCGGCAACGCTGTTG
TTGGGATGACGGACCTCGCTTCCACCTTATTCAGAACTTAAAGAGAAAT

>BigHornSD (132748)
TTCCAGCTCCAATAGTGTATATTAAAGTTGCTGTGATTAAAACGCTCGTAGTTGAATTA
TAAAGCGCTTAGCTTGCGGGTGCCCTGCTCTGCGGGGAGGAAGCTTGCGATAAACGGCGC
GGAGGCGATGCCGGCTTCGGCCGGTGTCATTACTTTGAAAAAATAGGGTGTTCAAAGCA
AATCTTATGTTAATGAATAATGAAGCATGGGACAATATCGAGGAGATCTTTCGGGATTT
CGGGAAAAGGATTAAGAGGAACAATTGGGGTGATTCAGAAAATGACGGGAGAGGTAAAA
TTCCATGATCGCTATAAGATGCACGAGAGCGAAAGCATTTCACTCAATTGCGTTCATTAA
TCAAGAACGAAAGTTAGGGGATCGAAGACGATCAGATACCGTCGTAGTCCTAACTATAA
ACGATGTCAACCAAGGATTGGATTAGTTTTAGAGAGACAGATGCTCGGTAACATCGTTA
TTGAGCGGACGGATCTCGCCTCCACCTTATTCAGAATTTAAAGAG

Parfrey, L. W., Walters, W. A., Lauber, C. L., Clemente, J. C., Berg-Lyons, D., Teiling, C., Kodira, C., Mohiuddin, M., Brunelle, J., Driscoll, M., Fierer, N., Gilbert, J. A. &

Knight, R. 2014. Communities of microbial eukaryotes in the mammalian gut within the context of environmental eukaryotic diversity. *Frontiers in Microbiology,* **5**.


Kanagroo1 (98428)
Sheep310 (FN666250)
Cow349 (FN666248)
Cow351 (FN666249)
Gazelle3 (6178)
Gazelle3 (4627)
Gazelle3 (47004)
Gazelle3 (136712)
Okapi2 (135300)
Okapi2 (45169)
Okapi2 (129418)
Okapi2 (107KBS209)
BigHornSD (132748)
Gazelle3 (93244)
BigHornSD (89533)
BigHornSD (78692)
BigHornSD (27883)
BigHornSD (42233)
BigHornSD (96034)
95/67/-
Okapi2 (85607)
Reindeer100 (FN666252)
Gazelle3 (7KBS3012)
BigHornSD (61415)
Sheep297 (FN666251)
Gazelle3 (42198)
Gazelle3 (102983)
Gazelle3 (113865)
Gazelle3 (43824)
Gazelle3 (53728)
E. bovis and variants
Okapi1 (17516)
92/*/1.0
Okapi1 (157076)
Okapi2 (39254)
Entamoeba CL4
99/86/ 1.0
Gazelle3 (37035)
Roedeer352 (FN666253)
BigHornSD (56084)
Entamoeba RL1
51/*/*
BigHornSD (28775)
99/82/ 0.99
BigHornSD (121866)
Okapi1 (6235)
Entamoeba CL5
Hulman (FR686358)
Entamoeba RL3
99/95/1.0
Cow349.2 (FR686362)
Cow350 (FR686363)
Entamoeba RL2
59/*/*
Gazelle3 (171795)
Gazelle3 (19865)
99/96/1.0
Okapi2 (3056)
Cow5 (KR025406)
Entamoeba RL8
Kangaroo1 (170614)
Entamoeba CL7
Okapi2 (92564)
Entamoeba CL3
Wildass1 (154451)
Entamoeba CL6
0.2

Entamoeba_RL10_(KR025408)																																				
E._hartmanni_(AF149907)	0.048																																			
E._ranarum_(AF149908)	0.382	0.376																																		
E._invadens_(AY769863)	0.520	0.510	0.221																																	
Entamoeba_RL7_(FR686360)	0.462	0.453	0.691	0.754																																
E._muris_(AB445018)	0.447	0.443	0.690	0.768	0.090																															
Entamoeba_RL11_(KR025409)	0.537	0.525	0.773	0.880	0.129	0.087																														
E._coli_ST1_(AF149915)	0.440	0.429	0.755	0.806	0.114	0.109	0.158																													
E._coli_ST2_(AF149914)	0.424	0.427	0.712	0.752	0.109	0.109	0.150	0.041																												
E._polecki_ST2_(AF149912)	0.447	0.452	0.697	0.795	0.538	0.506	0.574	0.534	0.526																											
E._polecki_ST4_(FR686357)	0.454	0.461	0.701	0.777	0.548	0.512	0.586	0.537	0.535	0.005																										
E._polecki_ST1_(AF149913)	0.475	0.497	0.705	0.795	0.527	0.498	0.559	0.514	0.512	0.015	0.017																									
E._polecki_ST3_(AJ566411)	0.463	0.491	0.740	0.821	0.581	0.546	0.589	0.563	0.565	0.020	0.020	0.018																								
E._suis_(DQ286372)	0.195	0.203	0.437	0.539	0.434	0.415	0.481	0.394	0.377	0.372	0.379	0.375	0.393																							
E._gingivalis_(D28490)	0.215	0.224	0.482	0.553	0.477	0.460	0.536	0.427	0.393	0.389	0.394	0.396	0.411	0.061																						
E._equi_(DQ286371)	0.161	0.161	0.355	0.492	0.437	0.424	0.490	0.405	0.396	0.428	0.428	0.444	0.463	0.201	0.225																					
E._terrapinae_(AF149910)	0.157	0.146	0.377	0.480	0.443	0.423	0.475	0.419	0.412	0.404	0.402	0.422	0.439	0.178	0.203	0.149																				
Entamoeba_RL2_(FR686363)	0.194	0.182	0.402	0.559	0.441	0.430	0.474	0.412	0.416	0.467	0.475	0.493	0.490	0.205	0.237	0.164	0.159																			
Entamoeba_RL2_(FR686362)	0.194	0.182	0.402	0.559	0.441	0.430	0.474	0.412	0.416	0.467	0.476	0.493	0.490	0.205	0.237	0.164	0.159	0.000																		
Entamoeba_RL3_(FR686359)	0.207	0.183	0.405	0.559	0.445	0.414	0.455	0.395	0.403	0.417	0.419	0.449	0.449	0.220	0.253	0.176	0.156	0.029	0.029																	
Entamoeba_RL3_(FR686358)	0.209	0.180	0.401	0.553	0.422	0.402	0.445	0.379	0.393	0.421	0.423	0.446	0.450	0.219	0.259	0.170	0.152	0.029	0.029	0.008																
Entamoeba_RL1_(FN666253)	0.200	0.191	0.421	0.547	0.471	0.416	0.458	0.394	0.387	0.447	0.449	0.478	0.474	0.225	0.257	0.161	0.160	0.035	0.035	0.032	0.031															
E._bovis_(FN666251)	0.204	0.182	0.428	0.557	0.445	0.424	0.461	0.389	0.412	0.448	0.456	0.494	0.493	0.210	0.242	0.161	0.153	0.033	0.033	0.035	0.031	0.032														
E._bovis_(FN666252)	0.199	0.179	0.414	0.536	0.437	0.421	0.458	0.386	0.400	0.442	0.450	0.488	0.487	0.207	0.247	0.157	0.147	0.033	0.033	0.032	0.030	0.031	0.008													
Entamoeba_RL8_(KR025406)	0.204	0.182	0.460	0.585	0.439	0.435	0.479	0.392	0.397	0.477	0.483	0.487	0.497	0.238	0.266	0.163	0.163	0.031	0.031	0.032	0.029	0.030	0.032	0.033												
Entamoeba_RL4_(FR686361)	0.164	0.160	0.390	0.552	0.420	0.392	0.437	0.372	0.362	0.450	0.459	0.465	0.479	0.177	0.224	0.137	0.107	0.089	0.089	0.094	0.090	0.091	0.085	0.083	0.096											
E._moshkovskii_(AF149906)	0.155	0.152	0.376	0.538	0.451	0.408	0.442	0.396	0.398	0.439	0.445	0.465	0.468	0.182	0.217	0.126	0.106	0.088	0.088	0.090	0.085	0.097	0.087	0.083	0.095	0.061										
E._ecuadoriensis_(DQ286373)	0.168	0.159	0.361	0.491	0.438	0.418	0.499	0.398	0.395	0.428	0.434	0.454	0.463	0.186	0.222	0.123	0.118	0.093	0.093	0.096	0.090	0.098	0.096	0.092	0.101	0.075	0.037									
E._bangladeshi_(KR025411)	0.164	0.155	0.364	0.488	0.434	0.414	0.493	0.391	0.380	0.435	0.440	0.460	0.466	0.188	0.221	0.121	0.119	0.090	0.090	0.091	0.084	0.095	0.093	0.089	0.097	0.073	0.032	0.005								
E._dispar_(Z49256)	0.161	0.157	0.374	0.493	0.429	0.425	0.496	0.398	0.390	0.441	0.440	0.472	0.466	0.186	0.228	0.133	0.110	0.093	0.093	0.094	0.087	0.099	0.094	0.090	0.101	0.074	0.033	0.014	0.011							
E._histolytica_(X56991)	0.157	0.150	0.377	0.497	0.421	0.416	0.488	0.398	0.390	0.429	0.434	0.462	0.463	0.179	0.218	0.130	0.104	0.089	0.089	0.090	0.084	0.095	0.090	0.086	0.096	0.068	0.032	0.013	0.012	0.005						
E._nuttalli_(FR686356)	0.159	0.152	0.379	0.500	0.423	0.418	0.491	0.400	0.392	0.431	0.437	0.465	0.465	0.181	0.220	0.131	0.105	0.090	0.090	0.092	0.085	0.096	0.091	0.087	0.097	0.069	0.033	0.013	0.011	0.004	0.001					
Entamoeba_RL6_(AF149911)	0.134	0.141	0.361	0.490	0.448	0.446	0.485	0.383	0.389	0.430	0.438	0.455	0.456	0.166	0.200	0.133	0.117	0.155	0.155	0.163	0.158	0.160	0.157	0.156	0.150	0.127	0.117	0.128	0.128	0.121	0.119	0.120				
Entamoeba_RL5_(FR686365)	0.141	0.135	0.373	0.471	0.468	0.448	0.530	0.429	0.422	0.424	0.435	0.458	0.458	0.187	0.206	0.120	0.117	0.146	0.146	0.149	0.151	0.148	0.148	0.143	0.152	0.110	0.108	0.113	0.107	0.109	0.106	0.107	0.070			
E._insolita_(AF149909)	0.133	0.122	0.350	0.460	0.449	0.443	0.507	0.414	0.411	0.412	0.423	0.441	0.444	0.174	0.213	0.113	0.103	0.126	0.126	0.131	0.129	0.127	0.123	0.120	0.127	0.105	0.100	0.104	0.100	0.107	0.100	0.101	0.076	0.054		
Entamoeba_CL2_(JQ406871)	0.147	0.178	0.375	0.500	0.498	0.469	0.569	0.408	0.420	0.465	0.442	0.501	0.448	0.294	0.230	0.084	0.022	0.150	0.150	0.166	0.157	0.170	0.168	0.168	0.159	0.096	0.079	0.091	0.098	0.098	0.091	0.098	0.116	0.071	0.084	
Entamoeba_RL9_(KR025407)	0.164	0.159	0.457	0.648	0.403	0.428	0.475	0.432	0.427	0.468	0.478	0.511	0.481	0.217	0.267	0.173	0.145	0.121	0.121	0.120	0.111	0.117	0.105	0.103	0.118	0.063	0.095	0.102	0.102	0.102	0.094	0.094	0.148	0.150	0.122	0.028