Chapter 2
The Continuously Expanding Universe of *Entamoeba*

C. Graham Clark and C. Rune Stensvold

Abstract In 1919, Clifford Dobell concluded that all the descriptions of *Entamoeba* in humans could be ascribed to three species: *Entamoeba histolytica*, *Entamoeba coli*, and *Entamoeba gingivalis*. At this time, morphology and host were the primary bases for naming species. We now know that both are unreliable, because host ranges can be broad and identical morphology can hide substantial genetic differences. Since Dobell, the number of accepted *Entamoeba* species in humans has continuously increased, with the most recent being identified in 2012. The application of molecular tools, especially DNA sequencing, has greatly increased our understanding of variation within the genus *Entamoeba*, but initial reliance on cultures gave us only a limited insight. For the past few years DNA extracted directly from feces from a wide range of hosts has been used to explore previously hidden *Entamoeba* diversity. Recent data include discovery of a uninucleate-cyst clade in nonhuman primates that is related to *Entamoeba bovis* and demonstration of substantial diversity within *E. coli*. Host ranges for some species are also expanding, with *E. coli* being found in rodents and an *E. muris*-like organism in primates. These results suggest that our picture of *Entamoeba* diversity is still incomplete and that further sampling is certain to uncover novel lineages.

2.1 Background

Amebae of the genus *Entamoeba* are easily recognized by their distinctive nucleus, which is usually described as being a "ring and dot" in appearance. The ring describes "peripheral chromatin" that lines the inside of the nuclear membrane, and the dot is generally referred to as the karyosome. The function of the karyosome

C.G. Clark (✉)
Department of Pathogen Molecular Biology, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, UK
e-mail: graham.clark@lshtm.ac.uk

C.R. Stensvold
Department of Microbiology and Infection Control, Statens Serum Institut, Copenhagen, Denmark

T. Nozaki, A. Bhattacharya (eds.), *Amebiasis*,
DOI 10.1007/978-4-431-55200-0_2

remains unknown, but in contrast to most eukaryotes this central body is not the nucleolus, as the functional equivalent of this organelle in *Entamoeba* is actually the peripheral chromatin [1]. With most stains this characteristic nuclear appearance is immediately recognizable, and an unidentified organism can be assigned to the genus *Entamoeba* on the basis of nuclear appearance alone.

It is the taxonomic ranks below that of genus that cause the majority of problems in the nomenclature and taxonomy of *Entamoeba*. The simplicity of the cell structure and lack of a fixed shape undermine the use of morphological features in species descriptions that is traditional in eukaryotic taxonomy. As a result, *Entamoeba* taxonomy has been subject to two types of "error" in the naming of species: (1) reliance on a character that does not reflect underlying genetic divergence, leading to overestimation of diversity and the naming of invalid species: examples include relying on the host as a species-specific character when in fact some *Entamoeba* species have quite a broad host range; and (2) morphological simplicity means that genetic divergence is not always reflected in morphological differences, which leads to underestimation of diversity and assigning the same species name to quite different organisms. Examples of both these errors follow.

Much of the interest in the taxonomy of *Entamoeba* arises from the need to positively identify *E. histolytica*, and the story of this species name is therefore a focus of the chapter. However, the taxonomy and nomenclature of the genus are currently in a state of flux, and this is also explored.

2.1.1 History of Entamoeba *Taxonomy from 1875 to 1919: Lösch to Dobell*

Although intestinal amebae had undoubtedly been seen by earlier workers, the first depiction of what is definitely *E. histolytica* is in the 1875 publication of Fedor Lösch [2], who was working in St. Petersburg, Russia. He describes the case of a peasant farmer, J. Markow, from the Archangel Territory of northern Russia, who was suffering from dysentery. Lösch's diagrams clearly show cells with the ring-and-dot nucleus characteristic of *Entamoeba*, and his description of the disease in Markow (and Lösch's animal model experiments) is fully compatible with invasive intestinal amebiasis. Lösch named the organism "Ameba coli" but whether he intended this as a descriptive name or a taxonomic name is unclear. The genus *Entamoeba* was named in 1895 by Casagrandi and Barbagallo [3], who were studying *E. coli*, and for reasons that are unclear they (and others) concluded that Lösch's ameba was the same organism. It fell to Fritz Schaudinn [4] to officially name *E. histolytica* in 1903 and distinguish it from the nonpathogenic *E. coli*, although he should actually have named the pathogen "*E. coli*" if he had followed the rules of nomenclature. Nevertheless, with a few exceptions, *E. histolytica* has been accepted as the name for the pathogenic species of *Entamoeba* in humans since that time.

By the time Dobell came to write his 1919 monograph "The amebae living in man" [5], the number of names for amebae that he considered to be wholly or partly synonyms of *E. histolytica* had risen to more than 30 and for *E. coli* to over 20. Some of these names were the result of disagreement over priority for the genus and species name, but many were the result of assigning species names to minor morphological variants of the amebae (sometimes imagined) or to amebae that were morphologically indistinguishable but from different hosts. Although we now know that Dobell made a number of wrong decisions, in his monograph he recognized only three species of *Entamoeba* in humans, namely *E. gingivalis* in the mouth plus *E. histolytica* and *E. coli* in the intestine. His strongly worded but well-reasoned case for this was highly influential in the field for many decades. By doing this "cull" of names he greatly simplified the nomenclature and placed the onus on future workers to justify why their "new" ameba should be considered a separate species by setting a standard against which others' descriptions should be compared.

2.1.2 *Taxonomy of* Entamoeba *in the Pre-molecular Era:* Entamoeba dispar, Entamoeba moshkovskii, Entamoeba hartmanni, *etc.*

The first "challenge" to Dobell's nomenclature was not long in arriving. In 1925 the great French parasitologist Emile Brumpt described *Entamoeba dispar* [6]. He concluded that this was a new species based on two observations. First, he observed an infection that persisted for many months without the infected individual developing symptoms. Second, he infected kittens, a very sensitive animal model for invasive amebiasis that was widely used at the time, and again no disease developed. He concluded that, despite the fact he could not distinguish them based on morphology or host, *E. dispar* and *E. histolytica* were distinct species [6]. No one believed him, or at least no one was willing to support this view publically. The Royal Society for Tropical Medicine and Hygiene invited Brumpt to give an address in 1926, and the record of the discussion that followed his presentation gives a fascinating insight into the general thinking at the time [7]. The issue was not so much whether *E. dispar* existed as a distinct species, but rather that it would be problematic because there was no way for them to be distinguished and therefore it would not inform medical decisions. For the next 50 years the species name virtually disappeared from the literature, only being used by Brumpt himself [8] and one of his students, Simic [9–13].

During the Second World War, a paper was published in Russian describing a new species of *Entamoeba* that was indistinguishable from *E. histolytica* by morphology but that had distinct physiological characteristics, namely, it was able to grow over a wide temperature range [14]. *Entamoeba moshkovskii* was isolated from sewage. Its isolation not from a host but from the environment raised questions

about whether it was a free-living organism or a parasite, an issue that still has not been fully resolved. Following the end of the war, environmental surveys across the world repeatedly isolated *E. moshkovskii* not only from sewage but also from freshwater sediments (see Clark and Diamond [15] for references). Clearly this organism is widespread if not ubiquitous. In the mid-1950s, an *Entamoeba* was isolated from a patient in Laredo, Texas, that was by chance found to grow at room temperature as well as at body temperature [16]. A few additional isolations of such amebae were made over the following years, but these "*E. histolytica*-like" ameba infections appeared to be rare (see Clark and Diamond [17] for references). The physiological similarity with *E. moshkovskii* did not escape notice, but proving they were the same organism was not possible in the absence of other criteria (see Sect. 2.2).

One of the synonyms of *E. histolytica* that Dobell had discarded in 1919 [5] was *Entamoeba hartmanni*. This species was described as closely resembling *E. histolytica* but significantly smaller: the cyst of *E. histolytica* is generally around 12 μm in diameter, while that of *E. hartmanni* is less than 10 μm, often around 7 μm, and the trophozoites are proportionately smaller also. Dobell believed that *E. hartmanni* represented one end of a size continuum and so was not distinct from *E. histolytica*. Nevertheless, "small race" *E. histolytica* was often referred to in the literature, indicating that other workers recognized it as an identifiable entity. In the 1950s, Burrows [18, 19] showed that the cyst size distribution in mixed infections showed two discrete peaks that coincided with the known mean sizes of the small and large "race" organisms. Therefore, although there was some overlap at the size extremes, "small race" *E. histolytica* was a stable and distinct entity. Burrows also noted some minor but consistent morphological differences in the nucleus to support his view and revived the name *E. hartmanni* for the small organism. This change was rapidly accepted in the field, suggesting that most workers already accepted that it was a distinct organism.

Thus, by the 1970s, *E. hartmanni* had joined *E. histolytica*, *E. coli*, and *E. gingivalis* as a member of the human *Entamoeba* flora, but the other proposed members had either been rejected (*E. dispar*) or were in limbo for lack of evidence (*E. moshkovskii*). The next steps required the development and application of new methodologies to the investigation of *Entamoeba*.

2.1.3 *The Early Use of Molecular but Non-sequence Data in Taxonomic Analysis: Isoenzymes, DNA Hybridization, and Antibody*

The first indication that Brumpt may have been correct about *E. dispar* being a distinct organism came in 1972. Martínez-Palomo and colleagues in Mexico used lectins to investigate the cell-surface sugars of *E. histolytica* growing in vitro. They observed that the effect of concanavilin A on agglutination of *E. histolytica*

divided strains into two groups and that these groups were correlated with whether the ameba had been isolated from an individual with disease or one without symptoms [20]: this was interpreted as revealing a virulence marker. A few years later, Sargeaunt and Williams started using the method of isoenzyme analysis, originally developed for and applied to bacteria, to investigate variation in *Entamoeba*. They initially showed that the method could distinguish between different species of *Entamoeba* [21] and uncovered intraspecific pattern variation. They later found that *E. histolytica* isolates fell into two groups (later called "pathogenic" and "nonpathogenic") that correlated with the disease status of the patient from whom the ameba was isolated [22]. They were later the first to revive Brumpt's proposal that two distinct species were involved. Soon thereafter, the first monoclonal antibodies against *E. histolytica* were developed; some of these also showed differential reactivity with strains depending on their origin [23]. Eventually, DNA analysis came on the scene, with Tannich et al. [24] showing that restriction fragment length polymorphisms (RFLP) revealed through Southern blotting also correlated with the isoenzyme patterns and strain origins. At about the same time, investigations using repetitive DNA of *E. histolytica* showed that some repeats only hybridized to one of the two isoenzyme groups [25].

2.2 The Impact of Ribosomal DNA on *Entamoeba* Taxonomy: Riboprinting, Sequences from Cultures, and Sequences from Stool DNA

Although evidence was accumulating in support of Brumpt's hypothesis, alternative explanations were still possible. For example, all the protein data could also be explained by differential expression of genes or different posttranslational modifications of proteins. The DNA data were trickier to explain, but differential gene amplification could not be ruled out, for example, if all strains had the same gene complement but one gene or genome "variant" was amplified in strains that caused disease. To investigate these alternative explanations, the ribosomal RNA (rRNA) genes were used. The small subunit (SSU) rRNA gene had emerged as a useful target for investigating phylogeny and molecular taxonomy because of its relatively slow evolutionary rate and high conservation within species. The SSU rRNA genes of "pathogenic" and "nonpathogenic" *E. histolytica* were first investigated in 1991 by Clark and Diamond [26] when it was shown that sequence differences detected using restriction enzymes (riboprinting) could again divide *E. histolytica* into two groups that correlated with the other markers reported previously. The basis of some of these restriction enzyme digestion differences at the sequence level was also investigated, and differential gene amplification using primer pairs that only amplified one of the two SSU rRNA gene variants found no evidence for the presence of the "other" variant: in other words, differential gene amplification could not be the explanation. Combined with accumulating sequences for other genes, by 1993 the

evidence was thought to be sufficient to warrant the redescription of *E. histolytica* to separate it from *E. dispar* and to recognize the existence of the latter species [27]. Brumpt was thus vindicated almost 70 years after his original proposal, and the existence of *E. dispar* appears to have been universally accepted. Nevertheless, the difficulty identified in 1926 [7] still exists: namely, differentiating the two species to inform medical decisions is still problematic. Microscopy is still the mainstay of parasite diagnosis in most parts of the world but does not allow the differentiation of *E. histolytica* and *E. dispar*.

Riboprinting also gave the first insight into the relationships among a wide range of species in the genus *Entamoeba*. This approach also showed the existence of cryptic diversity within several species, most notably *E. coli* and *E. moshkovskii* [15], and finally forged a link between *E. moshkovskii* and the rare "*E. histolytica*-like" infections represented by the Laredo isolate [17]. Crude estimates of similarity can be generated from the proportions of shared restriction sites, and these in turn can be used to generate phylogenetic trees. The vast majority of relationships identified using riboprinting estimates have been confirmed subsequently through sequencing of the SSU rRNA genes and the use of increasingly sophisticated phylogenetic analyses [28].

For many years the investigation of *Entamoeba* at the DNA level was limited to those species for which organisms growing in culture were available. This limitation also restricted the number of isolates of each that could be investigated as establishing an organism in culture is labor-intensive, and expensive if the culture is maintained for any length of time. Limitation of investigations to cultured organisms was in part a reflection of the difficulty of extracting usable DNA from fecal specimens. Feces is notorious for the presence of inhibitors of DNA analysis enzymes, and it was not until specific commercial kits were developed for fecal DNA extraction, eliminating this as a problem, that a full exploration of the genetic diversity of *Entamoeba* became possible.

That reliance on culture was possibly limiting our view of *Entamoeba* diversity was suspected early on, as an inability to grow certain species commonly found in livestock was noted. Whether the uninucleate cysts of *E. bovis* seen in cattle were genetically distinct from those of *E. polecki* seen in pigs, for example, could not be addressed because the former would not grow in culture. The advent of fecal DNA kits changed this, and it soon became clear that our view of *Entamoeba* diversity based on cultured amebae was woefully incomplete. The first SSU rRNA gene obtained from fecal DNA was from Vietnamese pigs and proved to be a new organism. Although expected to be *E. polecki* because of its uninucleate cysts, it proved to be unrelated and was given the name *E. suis* [29]. This finding was just the beginning, however: as new hosts are sampled, the known diversity of *Entamoeba* is still expanding rapidly. Even when more samples from previously investigated hosts are analyzed, it is not uncommon to detect new and genetically distinct sequence variants and even new branches of the *Entamoeba* tree. We have still only scratched the surface.

2.3 The Current Situation: Phylogeny, Diversity, Nomenclature, Host Specificity, and Impact of the Lack of Morphology

2.3.1 Recent Surge in Reports of Novel Ribosomal Lineages

The recent application of sequencing to polymerase chain reaction (PCR) products amplified from DNA extracted directly from feces has resulted in the discovery of a large number of novel *Entamoeba* lineages; to date, 29 distinct lineages have been identified (Table 2.1; Fig. 2.1) [30]. Some of these organisms were named a long time ago, while others appear to be distinct genetic variants (subtypes, ST) of species that have already been named (e.g., *E. coli* ST1 and ST2); however, further sampling may challenge this view—it is possible that the discreteness of lineages may disappear as further sequences become available. In addition, there are seven well-supported ribosomal lineages in the most recent phylogenetic trees that do not show a strong affinity with established species; these have been named by allocating a number to each ribosomal lineage (RL) (Fig. 2.1; Table 2.1) rather than a traditional Latin binomial. This terminology was introduced by Stensvold et al. [30] to enable a working nomenclature for *Entamoeba* organisms in the absence of morphological data and to respect the fact that names applicable to these lineages may already exist, but no data are available to link the two.

The recent and continuing surge in the discovery of new RLs and STs has made it clear that substantial additional sampling is needed to analyze the host specificity and genetic diversity of each potentially new lineage, and hence establishment of species names for novel lineages is still premature. An exception to this has been the recent finding of *Entamoeba bangladeshi* [36]. A variety of primers of varying specificity were applied to amplify DNA of *Entamoeba* directly from feces [30, 36], among them a genus-specific primer pair, ENTAGEN-F and ENTAGEN-R. This primer pair was used to analyze fecal DNAs from Bangladeshi children with and without diarrhea who were microscopy positive for four-nucleated cysts but PCR negative for *E. histolytica*, *E. dispar*, and *E. moshkovskii* [36]. Sequencing of PCR products produced evidence of a novel species, named *E. bangladeshi*. The morphology of cysts and trophozoite stages of *E. bangladeshi* appear similar to that of *E. histolytica* [36]. Phylogenetic analysis of the relationship between *E. bangladeshi* and other *Entamoeba* parasites reveals that, although distinct, *E. bangladeshi* clearly groups with the clade of *Entamoeba* infecting humans that includes *E. histolytica* and *E. dispar* (Fig. 2.1). *E. bangladeshi* is more distantly related than *E. dispar*, but closer than *E. moshkovskii,* to *E. histolytica* [36]: our more recent analyses suggest that it is specifically related to (but distinct from) *E. ecuadoriensis*, previously isolated only once, from sewage (unpublished observations). Hence, humans are now known to be hosts of four different but related species of *Entamoeba* (*E. histolytica, E. dispar, E. moshkovskii*, and *E. bangladeshi*), at least one of which is pathogenic, that cannot be differentiated using cyst morphology. Phylogenetically, this "complex" moreover contains an additional two species that have not been found in

Table 2.1 Host specificity of *Entamoeba*

Species/lineage	Subtype	Potentially invasive	Cyst nuclei[a]	Environment	Humans	Nonhuman primates	Ungulates	Rodents	Birds	Reptiles	Amphibia
E. histolytica	–	X	4	–	X	–	–	–	–	–	–
E. dispar	–	–	4	–	X	X	–	–	–	–	–
E. bangladeshi	–	–	4	–	X	–	–	–	–	–	–
E. moshkovskii	–	–	4	X	X	–	X[a]	–	–	X	–
E. nuttalli	–	X	4	–	–	X	–	–	–	–	–
E. ecuadoriensis	–	–	4	X	–	–	–	–	–	–	–
E. bovis	–	–	1	–	–	–	X	–	–	–	–
Entamoeba RL1	–	–	1	–	–	–	X	–	–	–	–
Entamoeba RL2	–	–	NA[b]	–	–	–	X	–	–	–	–
Entamoeba RL3	–	–	1	–	–	X	–	–	–	–	–
Entamobea RL4	–	–	NA	–	–	–	X	–	–	–	–
Entamoeba RL5	–	–	4	–	–	–	–	–	–	X	–
Entamoeba RL6	–	–	4	–	–	–	–	–	–	X	–
E. terrapinae	–	–	4	–	–	–	–	–	–	X	–
E. insolita	–	–	4	–	–	–	–	–	–	X	–
E. hartmanni	–	–	4	–	X	X	–	–	–	–	–
E. equi	–	–	NA	–	–	–	X	–	–	–	–
E. ranarum	–	–	1	–	–	–	–	–	–	–	X
E. invadens	–	X	1	–	–	–	–	–	–	X	–
E. suis	–	–	1	–	–	X	X	–	–	–	–
E. gingivalis	–	–	None	–	X	–	–	–	–	–	–

E. polecki	ST1	–	1	–	X	–	X	–	X	–	–
	ST2	–	1	–	X	X	–	–	–	–	–
	ST3	–	1	–	X	–	X	–	X	–	–
	ST4	–	1	–	X	–	–	–	–	–	–
Entamoeba RL7	–	–	8[c]	–	X[a]	X	X[d]	–	–	–	–
E. muris	–	–	8	–	–	–	–	X	–	–	–
E. coli	ST1	–	8	–	X	X	–	–	–	–	–
	ST2	–	8	–	X	X	–	X	–	–	–

For each species, subtype (ST) and ribosomal lineage (RL) the hosts recorded to data are indicated (X)
Based on references [29–35] and unpublished observations
[a]Unpublished observations
[b]Possibly four nuclei per cyst [30]
[c]Probably eight nuclei per cyst (Vidal-Lapiedra, unpublished observations)
[d]Lebbad et al., unpublished observations (1,306-bp sequence with 99 % identity to RL7)
NA - not available

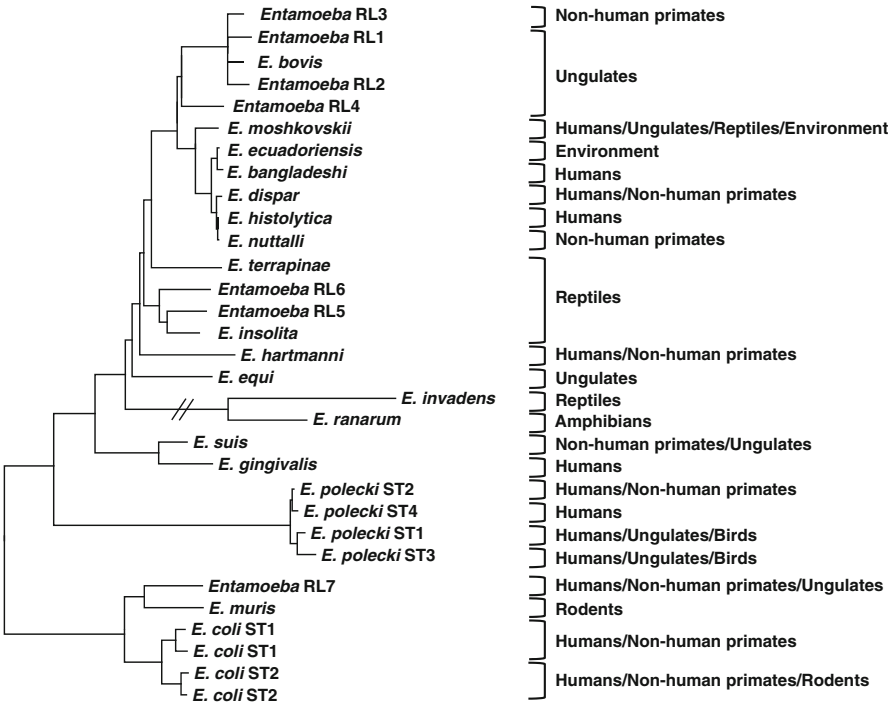

Fig. 2.1 Phylogenetic relationships among SSU rRNA gene sequences of *Entamoeba* species. The tree shown has been redrawn from Fig. 2.1 in Stensvold et al. [30], with the addition of *E. bangladeshi*. The cross-hatched branch has been shortened for convenience

humans so far, namely, *E. ecuadoriensis* (sewage) and *E. nuttalli* (nonhuman primates) [29, 31].

Designing PCR primers specific for a genus that keeps expanding in genetic diversity is challenging. Amplification of novel ribosomal lineages from fecal DNAs has been approached using the primers ENTAM1 and ENTAM2 [37], and more recently ENTAGEN-F and ENTAGEN-R [30]. The ENTAGEN primers amplify a PCR product of approximately 400 bp, and from the sequence of this product lineage-specific additional primers can be designed. An alternative approach is to apply low-specificity primers targeting all eukaryotic SSU rRNA genes, such as the RD5 + RD3 primer pair, with subsequent product cloning and clone sequencing; although this is even more sensitive, at least in theory, it can pose practical problems, such as nonspecific primer binding and the need for the screening of many clones that do not contain *Entamoeba* sequences to find those that do.

When screening for *Entamoeba*-specific DNA by PCR using genomic DNA extracted directly from feces, it appears, perhaps not surprisingly, that the choice of primers influences the results obtained in cases of mixed *Entamoeba* infection, as was recently observed during the screening of two samples from cattle [30].

Initially, unambiguous *E. bovis* SSU rRNA gene sequences were obtained using broad-specificity primers. Later, the application of primers of higher specificity led to the amplification and sequencing of a distinct lineage, RL2, from these two samples. Importantly, morphological data had been obtained before the molecular work, and although most of the cysts in the samples were uninucleate (consistent with *E. bovis*), a few tetranucleate cysts were seen as well; it is speculated that the tetranucleated cysts might represent the RL2 lineage. Hence, using primers of stricter specificity may be of great utility in cases where mixed infections are suspected.

Meanwhile, when designing diagnostic primers for a species the primers obviously need to take into account any intraspecies diversity that exists. There is evidence that published primers used for the detection of *E. hartmanni* target regions that exhibit sequence variation within this species, and that only some genetic variants of this species are being detected using that particular primer pair [30, 32]. A similar problem may exist in *E. moshkovskii*, where diagnostic primers have been designed based on the sequence of a single variant when many are known to exist.

2.3.2 Phylogeny and Host Specificity

Major host groups of *Entamoeba* identified to date include human and nonhuman primates, ungulates, rodents, reptiles, birds, and amphibians (Table 2.1; Fig. 2.1). An important observation is the fact that many lineages of great genetic diversity have been isolated from members of all host groups apart from birds and amphibians.

Entamoeba phylogeny indicates that the genus has not exclusively coevolved with its host but instead the parasite has repeatedly jumped host species barriers: how often is difficult to estimate. This variation is exemplified by the fact that *Entamoeba* lineages from, for example, humans and ungulates are scattered across the entire phylogenetic tree. Additionally, within the group of "ungulate-specific" lineages (in the topmost part of the tree in Fig. 2.1) is embedded RL3, which so far has only been found in various species of langur, arboreal monkeys from Asia. Moreover, although all four subtypes of *E. polecki* have been found in humans, only *E. polecki* ST2 (previously "*E. chattoni*") has been found in nonhuman primates; whether this is the result of insufficient sampling remains to be seen. Also of interest is the detection of a lineage related to *E. muris* (RL7) in a langur, *E. suis* in a gorilla, and *E. coli* in a chinchilla [30]. These observations taken together indicate that the host specificity of several lineages of *Entamoeba* may be only moderate, and some examples suggest a need to revisit and revise species' taxonomic descriptions.

Humans are natural hosts to at least eight well-established, genetically extremely diverse species of *Entamoeba*, and possibly more than 13 separate lineages (Table 2.1). Despite limited sampling, the relatively recent introduction of the use of PCR-based detection of *Entamoeba* directly from stool DNA has led to a total of 11 lineages being found in ungulates so far (Table 2.1). Most potential hosts remain to be investigated, and so it seems likely that the genetic universe of *Entamoeba* will continue expanding for the foreseeable future.

2.3.3 *Intraspecies Diversity of* Entamoeba

Studies of genetic diversity within *Entamoeba* lineages are pivotal for identification of differences in transmission patterns, including host specificity, evolutionary and taxonomic inferences, and for efforts to predict virulence and design relevant nucleic acid-based diagnostic tools. Intraspecies, and intra-RL, genetic diversity has been studied using SSU rRNA genes [15, 29, 30, 37–39], but multilocus markers are really needed for efforts to obtain high-resolution data (genotyping). The latter are starting to emerge and are addressed in another chapter of this book.

So far, genetic variation in species of *Entamoeba* other than *E. histolytica*, *E. dispar*, and *E. moshkovskii* has been little investigated, and this has consisted mainly of SSU rRNA gene analyses of other species of *Entamoeba* infecting humans.

E. polecki comprises four subtypes, that is, four clearly independent clusters within this well-defined species. All four subtypes have been found in humans, and humans are the only known host of ST4, which also appears to be the most common human *E. polecki* subtype. The host specificity of *E. polecki* subtypes is shown in Table 2.1. Interestingly, a mixed *E. polecki* infection was found for the first time in a sample from a rhea; the bird harbored both ST1 and ST3 [30]. *E. polecki* ST2 was formerly known as *E. chattoni* and *E. polecki* ST3 as *E. struthionis* [29, 30, 37, 40].

Similarly, *E. coli* has been divided into two STs based on phylogenetic analysis. ST1 has been detected only in samples from humans and nonhuman primates, whereas ST2 has also been found in one rodent so far (Table 2.1; [30]). Other distinct lineages are also characterized by octonucleate cysts, namely *E. muris* [41] and *Entamoeba* RL7 (A. Vidal-Lapiedra, unpublished observations). Here again it is not possible to ascribe species names to the RL, because there are a number of octonucleate cyst-producing taxa described for which no sequence data are available, including *E. caviae*, *E. cuniculi*, *E. gallinarum*, and *E. wenyoni* [41, 42], and the host spectrum of these lineages is still to be resolved; indeed, the possibility exists that these may be synonyms of *E. muris* or *E. coli*.

Quadrinucleate cyst-producing species of *Entamoeba* such as *E. hartmanni*, *E. dispar*, and *E. histolytica* appear to exhibit much less genetic variability [30, 43], except for *E. moshkovskii* ([15]; unpublished observations). Parija and Khairnar [43] studied the diversity in SSU rRNA genes of *E. histolytica*, *E. dispar*, and *E. moshkovskii* detected by nested multiplex PCR from various types of patient samples in Puducherry, India. Screening for polymorphisms employed riboprinting and single-strand conformation polymorphism (SSCP), and nucleotide sequencing was used for confirmation and identification of polymorphisms. A substantial number of isolates were screened, but RFLP analyses detected variation in only one *E. histolytica* isolate, suggesting a low degree of diversity in the SSU rRNA gene region studied. By SSCP, polymorphism was found in 7–12 % of the *E. histolytica* and *E. moshkovskii* PCR products, whereas no variation was found in the 174-bp SSU rRNA gene region of 171 *E. dispar* samples studied.

Substantial variation in *E. moshkovskii* was identified by riboprinting studies carried out by Clark and Diamond [15], who identified at least six ribodemes among 25 isolates; in situations where nucleic acid-based methods are used for primary detection (cysts of *E. moshkovskii* have only rarely been found in humans), problems with primer specificity may have led to detection of only certain variants of the species as mentioned earlier (Sect. 2.3.1; [44]).

Although the genetic diversity across the *E. dispar* SSU rRNA gene is probably <1 %, sequences from *E. coli* (including both subtypes) on the other hand exhibit 18 % polymorphic positions [45]; these estimates were based on data available in GenBank in 2010. It is worth noting that *E. coli* appears to exist in different "strains" distinguished by cyst size [46]; no attempt has yet been made to link this morphological difference to the SSU rRNA gene-based subtypes.

Few attempts have yet been made to compare variation in SSU rRNA genes to variation in protein-coding genes. The internal transcribed spacer (ITS) region, widely used to detect diversity in some organisms, has been studied only to a very limited extent in *E. histolytica* [47, 48], but data suggest that variation is substantially lower than in genes encoding repeat-containing proteins.

2.3.4 Impact of the Lack of Morphology

Detection, differentiation, and naming of species of *Entamoeba* based on morphological data are affected by at least three major problems.

1. For those ribosomal lineages only recently been discovered, very little information on morphology, if any, is available. Although this is often simply because of a lack of microscopic study, there have been cases in which *Entamoeba* sequences have been obtained from samples that were microscopy negative [30]. Traditionally, cysts have been isolated from fecal samples and analyzed by a number of techniques to produce the morphological description (see following). Although at least one species of *Entamoeba*, *E. gingivalis*, does not produce a cyst stage, the life cycles of many novel lineages are incompletely known. For instance, there are at least two different *Entamoeba* lineages infecting horses (unpublished data), and for neither of these has a cyst stage been identified despite careful microscopy having been performed.
2. Many infections involve multiple *Entamoeba* species, making the link between morphology and ribosomal sequences tenuous even when microscopy has been performed, as mentioned earlier for RL2 (Sect. 2.3.1), for example.
3. Apart from size and the number and appearance of nuclei and chromatoid bars, there are few morphological hallmarks that differentiate cysts, and the lack of features leads to two potential problems as mentioned in the earlier "Background" section. If we use minor morphological differences between cysts to separate species, these may turn out not to be reliable. Conversely, we may group together

distantly related organisms because we cannot tell them apart. Thus, we can either underestimate or overestimate diversity quite easily unless we supplement our morphological findings with molecular data. One such example is represented by the finding of uninucleate cysts in fecal samples from nonhuman primates. Such cysts have traditionally and collectively been referred to as "*Entamoeba chattoni*" (now *E. polecki* ST2), but recent sequencing of PCR products obtained from certain primate fecal samples containing only uninucleate cysts revealed the presence of a novel ribosomal lineage, *Entamoeba* RL3 [30].

Phylogenetic relationships observed between ribosomal lineages of *Entamoeba* analyzed by riboprinting [15] and, later, sequencing of cultured isolates [28, 29], until recently appeared to directly reflect the number of nuclei present in mature cysts. However, after the discovery of additional ribosomal lineages, primarily from mammals [30, 45], we now know that although lineages within most clades share the same number of nuclei in mature cysts, cyst nuclear number alone is not an indicator of phylogenetic relatedness among clades.

Nevertheless, detailed morphological analysis of trophozoites and cysts is still important. Erythrophagocytosis has been observed in a number of species, including *E. histolytica*, *E. invadens*, and *E. moshkovskii* [49–51], and is best visualised by permanent staining of fixed fecal smears. DAPI staining is one way of visualizing DNA and therefore enables counting of nuclei, and Calcofluor can be used for identification of a cyst wall in cases where it is necessary to differentiate trophozoites from cysts [38]. Immunofluorescence using a monoclonal antibody known to react with antigens in cysts of *E. histolytica* and *E. dispar*, but not *E. hartmanni*, *E. coli*, and *E. polecki*, remains to be tested on a wider range of lineages, and the significance of the specificity of the antibody is incompletely understood [38]. Conventional microscopy using iodine staining is helpful in the general morphological examination of cysts. The example taken from screening cattle samples for *Entamoeba* (see Sect. 2.3.1) emphasizes that thorough purification, staining, and morphological description is essential for cases in which mixed infections are present, to provide data critical to the interpretation of molecular data.

2.4 The Future

The identification of new *Entamoeba* lineages is likely to continue by "traditional" molecular methods as described here. However, a new source of data is likely to emerge in the near future, namely, the identification of *Entamoeba* by chance in environmental and fecal samples as part of metagenomic analyses and massive parallel sequencing of PCR products amplified by broad-specificity primers targeting the eukaryotic SSU rRNA gene. At present this is rare, because the approach used in most studies specifically targets bacterial ribosomal genes, but it seems likely that as costs decrease and applications broaden more and more eukaryote-oriented studies will be performed. We look forward with great interest to the new insights into *Entamoeba* diversity that such studies will provide!

References

1. Clark CG, Espinosa Cantellano M, Bhattacharya A (2000) *Entamoeba histolytica*: an overview of the biology of the organism. In: Ravdin JI (ed) Amebiasis. Imperial College Press, London
2. Lösch F (1875) Massenhafte Entwicklung von Amöben im Dickdarm. Virchow Arch 65:196–211
3. Casagrandi O, Barbagallo P (1895) Ricerche biologiche e cliniche suli' *Amoeba coli* (Lösch) (Nota preliminare). Bull Accad Gioenia Sci Nat Catania 39:4
4. Schaudinn F (1903) Untersuchungen über die Fortpflanzung einiger Rhizopoden (Vorläufige Mittheilung). Arbeit Kaiserlichen Gesundheitsamte 19:547–576
5. Dobell C (1919) The amoebae living in man, A zoological monograph. J. Bale, Sons, and Danielson, London
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 115–124)
8. Brumpt E (1936) Précis de parasitologie. Masson et Cie, Paris
9. Simic T (1931) Présence de l'*Entamoeba dispar* Brumpt dans le sud de la Yougoslavie (Macédoine Serbe). Ann Parasitol Hum Comp 9:289–302
10. Simic T (1931) Infection expérimentale de l'homme par *Entamoeba dispar* Brumpt. Ann Parasitol Hum Comp 9:385–391
11. Simic T (1931) Étude expérimentale complémentaire de l'*Entamoeba dispar* Brumpt, de Skoplje, sur le chat. Ann Parasitol Hum Compar 9:497–502
12. Simic T (1935) Infection expérimentale du chat et du chien par *Entamoeba dispar* et *Entamoeba dysenteriae*. Réinfection et immunité croisée du chien. Ann Parasitol Hum Comp 13:345–350
13. Simitch T, Petrovitch Z (1953) Recherche de protozoaires intestinaux chez l'homme par culture a partir de selles fraiches. Ann Parasitol Hum Comp 28:157–162
14. Tshalaia LE (1941) On a species of *Entamoeba* detected in sewage effluents (in Russian). Med Parazit (Moscow) 10:244–252
15. Clark CG, Diamond LS (1997) Intraspecific variation and phylogenetic relationships in the genus *Entamoeba* as revealed by riboprinting. J Eukaryot Microbiol 44:142–154
16. Dreyer DA (1961) Growth of a strain of *Entamoeba histolytica* at room temperature. Tex Rep Biol Med 19:393–396
17. Clark CG, Diamond LS (1991) The Laredo strain and other *Entamoeba histolytica*-like amoebae are *Entamoeba moshkovskii*. Mol Biochem Parasitol 46:11–18
18. Burrows RB (1957) *Endamoeba hartmanni*. Am J Hyg 65:172–188
19. Burrows RB (1959) Morphological differentiation of *Entamoeba hartmanni* and *E. polecki* from *E. histolytica*. Am J Trop Med Hyg 8:583–589
20. Martínez-Palomo A, González-Robles A, De la Torre M (1973) Selective agglutination of pathogenic strains of *Entamoeba histolytica* induced by Con A. Nat New Biol 245:186–187
21. Sargeaunt PG, Williams JE (1978) Electrophoretic isoenzyme patterns of *Entamoeba histolytica* and *Entamoeba coli*. Trans R Soc Trop Med Hyg 72:164–166
22. Sargeaunt PG, Williams JE, Grene JD (1978) The differentiation of invasive and non-invasive *Entamoeba histolytica* by isoenzyme electrophoresis. Trans R Soc Trop Med Hyg 72:519–521
23. Strachan WD, Spice WM, Chiodini PL, Moody AH, Ackers JP (1988) Immunological differentiation of pathogenic and non-pathogenic isolates of *Entamoeba histolytica*. Lancet 1:561–563
24. Tannich E, Horstmann RD, Knobloch J, Arnold HH (1989) Genomic DNA differences between pathogenic and nonpathogenic *Entamoeba histolytica*. Proc Natl Acad Sci USA 86: 5118–5122
25. Bracha R, Diamond LS, Ackers JP, Burchard GD, Mirelman D (1990) Differentiation of clinical isolates of *Entamoeba histolytica* by using specific DNA probes. J Clin Microbiol 28:680–684

26. Clark CG, Diamond LS (1991) Ribosomal RNA genes of 'pathogenic' and 'nonpathogenic' *Entamoeba histolytica* are distinct. Mol Biochem Parasitol 47:297–302
27. Diamond LS, Clark CG (1993) A redescription of *Entamoeba histolytica* Schaudinn, 1903 (emended Walker, 1911) separating it from *Entamoeba dispar* Brumpt, 1925. J Eukaryot Microbiol 40:340–344
28. Silberman JD, Clark CG, Diamond LS, Sogin ML (1999) Phylogeny of the genera *Entamoeba* and *Endolimax* as deduced from small-subunit ribosomal RNA sequences. Mol Biol Evol 16:1740–1751
29. Clark CG, Kaffashian F, Tawari B, Windsor JJ, Twigg-Flesner A, Davies-Morel MCG, Blessmann J, Ebert F, Pesche B, Le Van A, Jackson CJ, 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
30. Stensvold CR, Lebbad M, Victory EL, Verweij JJ, Tannich E, Alfellani M, Legarraga P, Clark CG (2011) Increased sampling reveals novel lineages of *Entamoeba*: consequences of genetic diversity and host specificity for taxonomy and molecular detection. Protist 162:525–541
31. Tachibana H, Yanagi T, Pandey K, Cheng XJ, Kobayashi S, Sherchand JB, Kanbara H (2007) An *Entamoeba* sp. strain isolated from rhesus monkey is virulent but genetically different from *Entamoeba histolytica*. Mol Biochem Parasitol 153:107–114
32. Suzuki J, Kobayashi S, Murata R, Tajima H, Hashizaki F, Yanagawa Y, Takeuchi T (2008) A survey of amoebic infections and differentiation of an *Entamoeba histolytica*-like variant (JSK2004) in nonhuman primates by a multiplex polymerase chain reaction. J Zoo Wildl Med 39:370–379
33. Levecke B, Dreesen L, Dorny P, Verweij JJ, Vercammen F, Casaert S, Vercruysse J, Geldhof P (2010) Molecular identification of *Entamoeba* spp. in captive nonhuman primates. J Clin Microbiol 48:2988–2990
34. Verweij JJ, Vermeer J, Brienen EA, Blotkamp C, Laeijendecker D, van Lieshout L, Polderman AM (2003) *Entamoeba histolytica* infections in captive primates. Parasitol Res 90:100–103
35. Feng M, Yang B, Yang L, Fu Y, Zhuang Y, Liang L, Xu Q, Cheng X, Tachibana H (2011) High prevalence of *Entamoeba* infections in captive long-tailed macaques in China. Parasitol Res 109:1093–1097
36. Royer TL, Gilchrist C, Kabir M, Arju T, Ralston KS, Haque R, Clark CG, Petri WA Jr (2012) *Entamoeba bangladeshi* nov. sp., Bangladesh. Emerg Infect Dis 18:1543–1545
37. Verweij JJ, Polderman AM, Clark CG (2001) Genetic variation among human isolates of uninucleated cyst-producing *Entamoeba* species. J Clin Microbiol 39:1644–1646
38. Stensvold CR, Lebbad M, Clark CG (2010) Genetic characterisation of uninucleated cyst-producing *Entamoeba* spp. from ruminants. Int J Parasitol 40:775–778
39. Takano J, Tachibana H, Kato M, Narita T, Yanagi T, Yasutomi Y, Fujimoto K (2009) DNA characterization of simian *Entamoeba histolytica*-like strains to differentiate them from *Entamoeba histolytica*. Parasitol Res 105:929–937
40. Martínez-Díaz RA, Herrera S, Castro A, Ponce F (2000) *Entamoeba* sp. (Sarcomastigophora: Endamoebidae) from ostriches (*Struthio camelus*) (Aves: Struthionidae). Vet Parasitol 92: 173–179
41. Neal RA (1950) An experimental study of *Entamoeba muris* (Grassi, 1879); its morphology, affinities and host–parasite relationship. Parasitology 40:343–365
42. Levine ND (1973) Protozoan parasites of domestic animals and man. Burgess, Minneapolis
43. Parija SC, Khairnar K (2008) Mutation detection analysis of a region of 16S-like ribosomal RNA gene of *Entamoeba histolytica*, *Entamoeba dispar* and *Entamoeba moshkovskii*. BMC Infect Dis 8:131
44. Stensvold CR, Lebbad M, Verweij JJ, Jespersgaard C, von Samson-Himmelstjerna G, Nielsen SS, Nielsen HV (2010) Identification and delineation of members of the *Entamoeba* complex by pyrosequencing. Mol Cell Probes 24:403–406
45. Stensvold CR, Lebbad M, Verweij JJ (2011) The impact of genetic diversity in protozoa on molecular diagnostics. Trends Parasitol 27:53–58

46. Matthews JR (1919) A mensurative study of the cysts of *Entamoeba coli*. Ann Trop Med Parasitol 12:259–272
47. Som I, Azam A, Bhattacharya A, Bhattacharya S (2000) Inter- and intra-strain variation in the 5.8S ribosomal RNA and internal transcribed spacer sequences of *Entamoeba histolytica* and comparison with *Entamoeba dispar*, *Entamoeba moshkovskii* and *Entamoeba invadens*. Int J Parasitol 30:723–728
48. Ghosh S, Frisardi M, Ramirez-Avila L, Descoteaux S, Sturm-Ramirez K, Newton-Sanchez OA, Santos-Preciado JI, Ganguly C, Lohia A, Reed S, Samuelson J (2000) Molecular epidemiology of *Entamoeba* spp.: evidence of a bottleneck (demographic sweep) and transcontinental spread of diploid parasites. J Clin Microbiol 38:3815–3821
49. Tsutsumi V, Ramírez-Rosales A, Lanz-Mendoza H, Shibayama M, Chávez B, Rangel-López E, Martínez-Palomo A (1992) *Entamoeba histolytica*: erythrophagocytosis, collagenolysis, and liver abscess production as virulence markers. Trans R Soc Trop Med Hyg 86:170–172
50. Ramírez-Córdova JD, Ramírez-Romo S, Mora-Galindo J (1990) Quantitative evaluation of intracellular degradation in *Entamoeba invadens*. Int J Parasitol 20:199–201
51. Pizzi T (1956) Observations on erythrocyte phagocytosis by *Entamoeba moshkovskii* as reported by Tshalaia, 1941 (in Spanish). Bol Chil Parasitol 11:7–9