Experimental Parasitology 145 (2014) S39–S45

Contents lists available at ScienceDirect

Experimental Parasitology

journal homepage: www.elsevier.com/locate/yexpr

Insights from the DNA databases: Approaches to the phylogenetic structure of *Acanthamoeba*

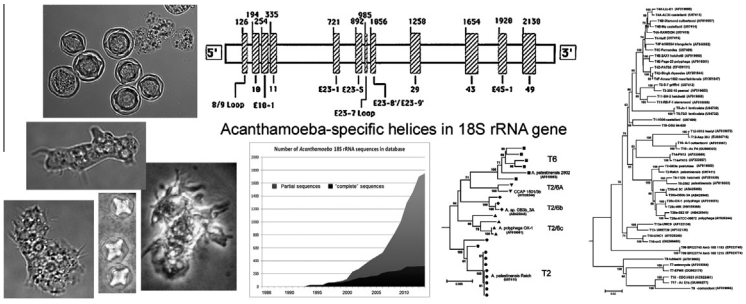
Paul A. Fuerst *

Department of Evolution, Ecology and Organismal Biology, The Ohio State University, 388 Aronoff Laboratory, 318 W. 12th Ave, Columbus, OH 43210, USA
Department of Molecular Genetics, The Ohio State University, 388 Aronoff Laboratory, 318 W. 12th Ave, Columbus, OH 43210, USA

HIGHLIGHTS

- *Acanthamoeba Rns* sequences in the international DNA databases increased to >1800 since 1986.
- Phylogenetic study of "complete" sequences refines sequence types introduced (1996) by Byers–Fuerst lab.
- Analysis relates *Rns* sequence types and morphological groups and species of *Acanthamoeba*.
- Sequence types clearly correlate with groups but not with nominal species of *Acanthamoeba*.
- Significant phylogenetic groups exist within sequence types, and may represent species.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 6 December 2013
Received in revised form 14 June 2014
Accepted 24 June 2014
Available online 3 July 2014

Keywords:
Acanthamoeba
18S rRNA gene
Phylogenetic analysis
Species names
Sequence types

ABSTRACT

Species of *Acanthamoeba* have been traditionally described using morphology (primarily cyst structure), or cytology of nuclear division (used by Pussard and Pons, 1977). Twenty-plus putative species were proposed based on such criteria. Morphology, however, is often plastic, dependent upon culture conditions. DNA sequences of the nuclear small subunit (18S) rRNA that can be used for the study of the phylogeny of *Acanthamoeba* have increased from a single sequence in 1986 to more than 1800 in 2013. Some of the patterns of the sequence data for *Acanthamoeba* are reviewed, and some of the insights that this data illuminates are illustrated. In particular, the data suggest the existence of 20 or more genotypic types, a number not dissimilar to the number of named species of *Acanthamoeba*. However, molecular studies make clear that the relationship between phylogenetic relatedness and species names as we know them for *Acanthamoeba* is tenuous at best.

1. Background

The taxonomic classification of free-living amoebae has always been challenging, given the pliable nature of amoebic cell size and shape. Since the early part of the last century, taxonomic classification of small free-living amoebas was based primarily on morphological criteria (Pussard and Pons, 1977; Visvesvara, 1991). These included the type of locomotion of the trophozoites, the morphology of the cysts and the type of nuclear division of the organism. By the late 1970s, however, new methods of analysis began to impact taxonomy. The categorization of species, especially for simple microbial organisms such as the free-living amoebae, began to come under scrutiny. Previously accepted morphological criteria

* Corresponding author at: Department of Evolution, Ecology and Organismal Biology, The Ohio State University, 388 Aronoff Laboratory, 318 W. 12th Ave, Columbus, OH 43210, USA. Fax: +1 614 292 2030.
E-mail address: fuerst.1@osu.edu

http://dx.doi.org/10.1016/j.exppara.2014.06.020
0014-4894/

began to be questioned as biochemical and, later, molecular criteria indicated inconsistencies in the patterns of classification. Biochemical approaches to systematic classification used similarity of allozyme patterns to group individuals and categorize species boundaries in *Acanthamoeba* on the basis of similar protein patterns (De Jonckheere, 1983; Daggett et al., 1982, 1985). They came to the conclusion that the taxonomic assignments used at that time did not, in general, correspond to biochemically distinguishable lineages. Biochemical groups usually included members assigned to multiple nominal species, and isolates assigned to the same nominal taxa were often found in different biochemical lineages.

2. The Introduction of DNA sequencing and the 18S rRNA gene

By 1980, DNA analysis had begun to replace isozymes as the methods of choice for evolutionary genetics. The seminal contribution that permitted DNA sequencing to be applied extensively was the report of the sequence of the nuclear small subunit (18S) rRNA gene for the Neff strain of *Acanthamoeba castellanii* (Gunderson and Sogin, 1986). The most important aspect reported in this paper, other than the fact that it represented the first 18S rRNA gene sequence from *Acanthamoeba*, was the fact that *A. castellanii* Neff possessed an 18S rRNA gene that had unusually long length and (possibly) unusual characteristics. The gene was almost 500 nucleotides longer than a typical eukaryotic 18S rRNA gene. Further, the increase in length was due to expansion of a series of regions that were dispersed within the gene.

It quickly became clear that the nuclear 18S rRNA genes (designated *Rns*) of *Acanthamoeba* contained much more information about interstrain relationships than was usually obtained when sequencing eukaryotic 18S rRNA genes. In most situations, the various strains of a particular genus are likely to have somewhat similar nucleotide sequences. Natural selection acts to maintain the functioning of the rRNA genes, because they are vital for the life of the cell. As a consequence, changes in rRNA gene sequences over evolutionary time are usually slow. Within-species polymorphisms in these genes has usually been found to restricted. It is not uncommon to find that most members of a species exhibit the same sequence. Observations in which different conspecific individuals of animals or plants differ by more than a few nucleotides (out of ~1800 nucleotides in a typical eukaryotic 18S rRNA) are rare. The definition of "species" in single-celled eukaryotes becomes more difficult, but the observation that low rRNA diversity is usually observed between "conspecific" organisms still generally holds. However, that did not seem to be the case for "species" of *Acanthamoeba*. It quickly became obvious that the extra sequence regions discovered by Gunderson and Sogin (1986) were probably subjected to much weaker purifying natural selection, and thus were able to vary much more extensively than the majority of nucleotide sites within the gene. These regions would provide the necessary information for an extensive analysis of phylogenetic relationships among isolates of *Acanthamoeba*.

Following the publication of the first full sequence of the 18S rRNA gene from "*A. castellanii* Neff" additional sequences began to be determined. An increase in the number of *Rns* sequences determined in the early 1990's allowed the first sequence based analysis of genotype diversity within *Acanthamoeba* to identify patterns of phylogenetic relationships. The first extensive paper on this approach was contributed by our lab (Gast et al., 1996), and proposed that these sequences could be used for subgenus classification. In our original paper we defined the sequence types on the basis of criteria that we felt provided significant separation of types while retaining isolates together that were substantially similar, but not identical. Our original definition was "*types are defined as sequences or groups of sequences that differ from all other sequences by at least 6%, have a minimum of 134 base differences, or an evolutionary distance greater than 0.8% in the current database*" (Gast et al., 1996). We never assumed that this definition would be set in stone, since the number of sequences remained quite small. We intended this definition to be applied to sequences that were substantially complete, with respect to the original sequence of Gunderson and Sogin (1986), and did not envision that it would be applied to sequence fragments that represented a small fraction of the total gene.

In the first definition of the sequence type, we defined four genotype clusters from among a collection of 18 isolates. These were designated type T1 through T4. The analysis included information from the original *Acanthamoeba* isolate of *A. castellanii* (ATCC 30011) and the isolate "*A. castellanii* Neff" (ATCC 30010). These two sequences fell within the same sequence type, T4. Type T4 was the predominant form that we observed and this has remained the case as the number of sequenced isolates expanded. In the initial set of 18 isolates, no sequence was identical to any other, but 15 of the 18 isolates were designated T4.

The paper of Gast et al. (1996) on genotype groups also identified an important aspect of sequence variation within the *Rns* sequences of *Acanthamoeba*. Twelve variable regions of the gene were identified, most of which corresponded to sequence regions that do not have obvious homologs in the *Rns* genes of other eukaryotes. Eventually many of these regions were identified as hypervariable sequence regions and have become the basis for further analysis tracking disease cases as well as attempting to further subdivide sequence types and endeavoring to bring together species identification with sequence genotype analysis. Within two years of our initial proposal of sequence types, by adding the sequences of an additional 35 isolates we had identified an additional 8 types within *Acanthamoeba* (Stothard et al., 1998). This study was important because it included representatives of all three morphological groups, and considered previous species designations by including isolates that represented at least 16 named species of *Acanthamoeba*.

In the Stothard et al. (1998) study we revised our definition of sequence types based on our ability to identify monophyletic lineages that roughly corresponded to the diversity we observed in the original sample of T4 isolates (Gast et al., 1996). We found that sequence differences between types were at least 5%, and were always greater than sequence differences within types. However, in that paper we never formally proposed the 5% difference criterion that has shaped considerations by ourselves and others of subsequent sequence discoveries. Nevertheless, the use of a 5% criterion for the definition of new sequence types has become entrenched in the analysis of *Acanthamoeba*. In preparing this paper, an analysis of more than 330 *Acanthamoeba* small subunit rRNA sequences exceeding 2000 bases in length suggests that, while the arbitrary 5% cutoff is not grossly out of line, a more appropriate value would be 4% (Fuerst et al., 2014). This would separate almost all isolates that belong to different significant monophyletic clades that are equivalent to our original sequence types. However, it must be emphasized that even this value should be used judiciously, especially in cases where a new proposed type is represented by only one or two isolates. Furthermore, when we examine the existing large dataset being reviewed here, using our original approach of identifying significant monophyletic lineages, we find support for the idea that we can identify formal sub-types within sequence types. In this case formal sequence sub-types would have similarities that distinguish each sub-type, but in which sub-types differ by much less than 5%. Analysis of differences between possible significant sub-types (to be reported elsewhere) suggests that different sub-types are usually characterized by a level of pairwise sequence divergence of greater than 2% but less than 4%. Application of this sub-type definition must include

careful examination to account for sequencing errors. Much additional details concerning sequence similarities cannot be provided because of space considerations. A series of papers detailing the phylogenetic relationships among isolates and relating sequence types to species names and levels of differentiation is being prepared (see Fuerst et al., 2014 for the first part of this analysis).

Since our last major compilation of sequence types (Stothard et al., 1998), a number of other investigators have reported sequences that do not fall into the 12 sequence types that we described earlier (Horn et al., 1999; Gast, 2001; Hewett et al., 2003; Lanocha et al., 2009; Corsaro and Venditti, 2010; Nuprasert et al., 2010; Qvarnstrom et al., 2013). There appear to be about 20 sequence types that can be currently identified. However, some confusion exists in the literature, in some cases with divergent sequences being assigned the same new type number (Lanocha et al., 2009; Corsaro and Venditti, 2010). There have also been some claims of new types that have accompanied only partial *Rns* sequences (Hewett et al., 2003; Lanocha et al., 2009). While these may ultimately be validated, we strongly recommend that no claim of a new type be made without information on a sequence that spans essentially the entire *Rns* gene. (In fact, we would propose a multigene analysis in the future.)

3. The identification of diagnostic fragments and their effect on isolate screening

One of the main goals that our lab had for the analysis of *Acanthamoeba* was to develop rapid screening procedures. While the ability of obtaining almost complete *Rns* sequences was an ideal goal, we examined whether smaller parts of the *Rns* sequence contained a substantial proportion of the signal from the entire gene. This would speed up the collection of sequence information, and possibly speed the ability of rapidly passing diagnostic information back to the clinician studying possible infections involving *Acanthamoeba*. Our goal was to identify portions of the *Rns* sequence that would be (i) highly specific for the genus *Acanthamoeba*, (ii) obtainable from all known genotypes, and (iii) useful for identification of individual genotypes. We were able to identify specific regions of the *Rns* sequence that met these goals, and which could be obtained easily by PCR (Schroeder et al., 2001). One PCR amplimer, originally designated ASA.S1 (*Acanthamoeba*-specific amplimer S1) and slightly greater than 400 nucleotides in length in most sequence types, satisfied requirements (i) and (ii) and resulted in a product that would be diagnostic for the presence of *Acanthamoeba*. This amplimer did not, however, appear to distinguish between all sequence types. A second set of PCR primers were identified that would provide genotype identification. We proposed that three amplimers should be routinely obtained and designated these diagnostic fragments DF1, DF2, and DF3 (Schroeder et al., 2001). It turned out that the most informative of these, DF3, which is a fragment of about 240 nucleotides, was included within the bounds of amplimer ASA.S1.

Identification of isolates using the diagnostic fragments associated with either ASA.S1 or DF3 has become the most widely utilized screening tool, as measured by reports of studies identifying *Acanthamoeba* in either clinical or environmental samples. Within the DNA databases, exemplified by GenBank, approximately 220 bulk submissions (DNA submissions of 1 or more sequences) involving either a publication or a proposed publication occurred through the end of 2013. Of these, 115 involved sequences obtained with the use of ASA.S1, 54 used DF3, while 63 included almost complete sequences. About twice as many sequences in the DNA databases have been obtained using ASA.S1 compared to only DF3.

4. The growth in the number of *Acanthamoeba* sequences in the international DNA databases

From the time of the first sequence report by Gunderson and Sogin (1986) to the present has seen a remarkable increase in the level of interest and in the number of reports from workers around the world concerning the genetic diversity of *Acanthamoeba* that have been found in clinical, nonclinical and environmental settings. We have endeavored to compile this information as it has been deposited in the international DNA databases, but have also included sequences that in the course of a study or clinical investigation have been collected but were not deposited for further use. Many researchers generously responded to our inquiries, providing us with information concerning such undeposited sequences. These have been collated together with information of

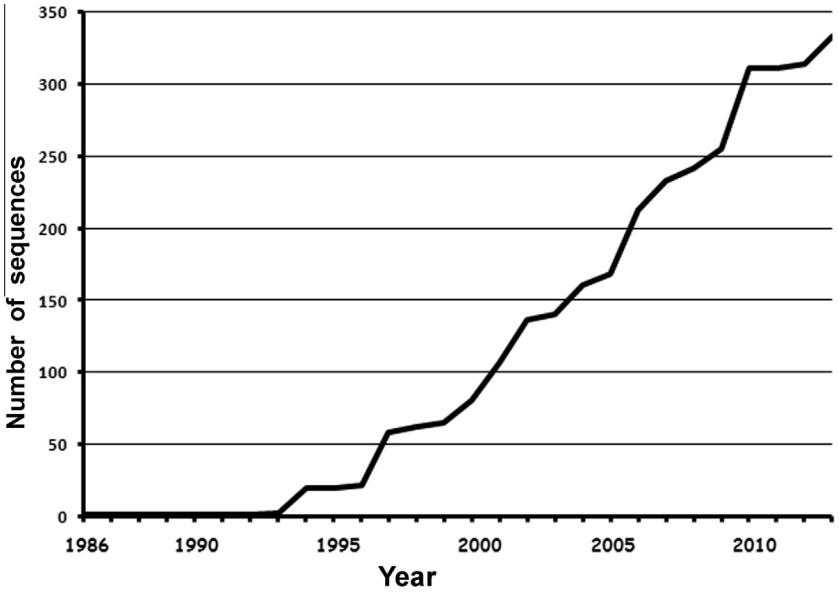


Fig. 1. Cumulative increase by year in the number of *Acanthamoeba Rns* sequences greater that 2000 nucleotides in length in the international DNA databases through November 2013.

sequences from the databases. We will be presenting detained information from all of our collations on an updated *Acanthamoeba* sequence database that is replacing an earlier version that had been maintained at The Ohio State University. Numbers that are presented in this report include information from both the sequences that have been deposited from other sequences that have been provided to us by our many generous collaborators.

Since 1986, the number of *Acanthamoeba Rns* sequences has increased substantially. In November 2013, data on nearly complete (sequences >2000 nucleotides in length) or partial *Rns* sequences had been reported for over 1820 isolates. This extensive set of data allows a substantial number of questions to be investigated, only a few of which can be presented here. The most phylogenetically informative subset of this data is represented by *Rns*

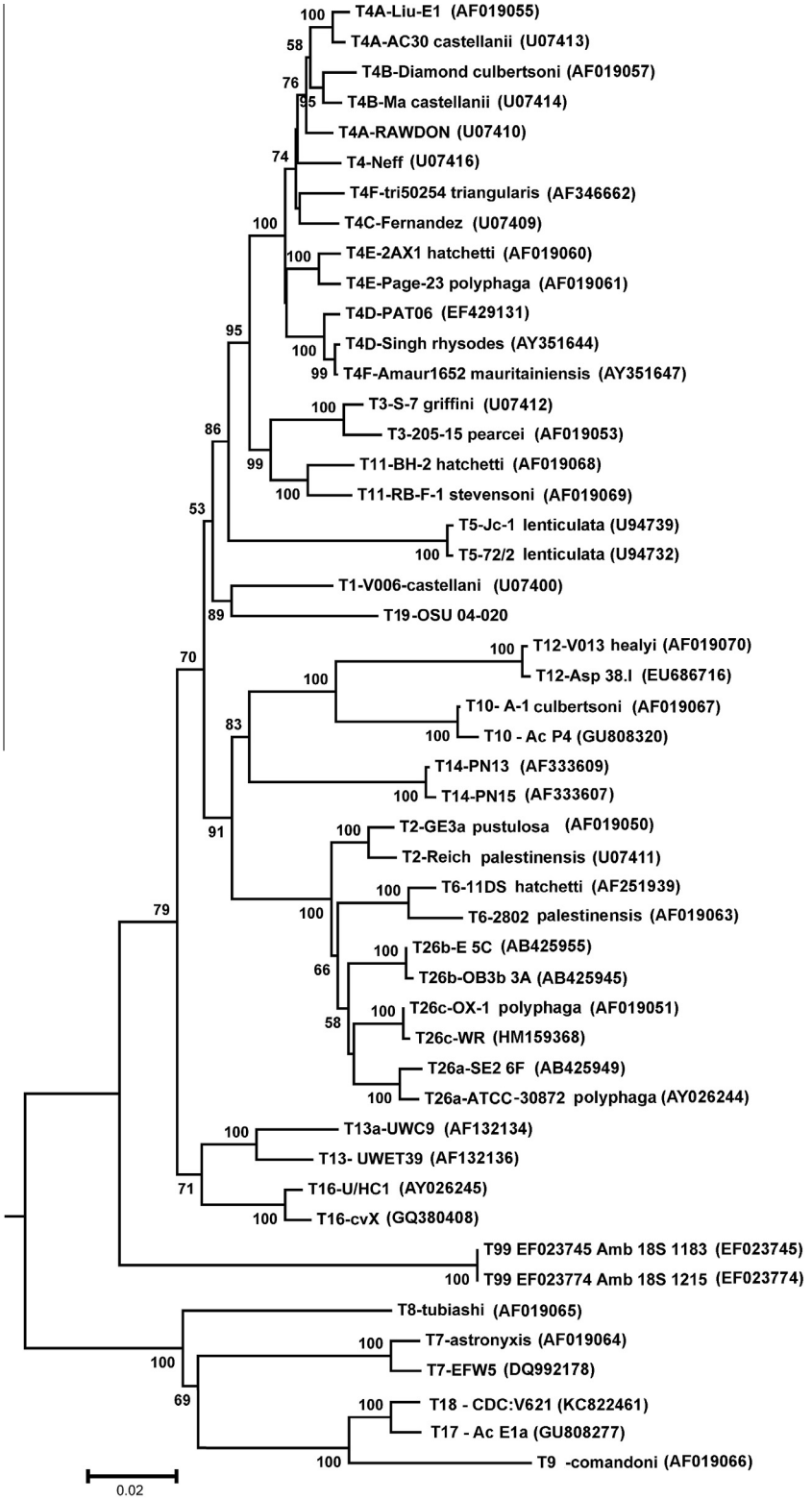


Fig. 2. Phylogenetic relationships among various representatives of each of the current sequence types as determined from *Acanthamoeba Rns* sequences greater than 2000 bases. The tree was constructed using the neighbor joining method.

sequences that are 2000 nucleotides or longer. Fig. 1 presents the yearly pattern of increase in the number of these "complete" *Rns* sequences in the DNA databases. By November 2013, 333 such sequences had been collected. Cumulative information on the growth of DNA sequences in the database and the assignment of types and subtypes to isolates is available at http://u.osu.edu/acanthamoeba/.

5. Distribution of isolates among standard genotypes

We have examined the phylogenetic relationship among these long sequences, but will here present only summaries of our findings, with the intention of presenting extensive details elsewhere. The phylogenetic relationships encoded within this set of sequences has been analyzed using several different methods, including neighbor joining (Saitou and Nei, 1987), Maximum Likelihood (developed following Felsenstein, 1981) and Bayesian analysis (Huelsenbeck and Ronquist, 2001). All methods resulted in the same general insights into the relationship among genotypes, which we will discuss here, although specific relationships between individual isolates were less likely to be the same when different methods were used. The general relationships between various sequence types within *Acanthamoeba* are shown in the phylogenetic tree in Fig. 2. This tree shows only a representative small subset of isolates from which the analysis was produced, in order to provide a general indication of the patterns of divergence. It does allow us to determine standard sequences that can be used to place partial sequences accurately within sequence types.

While sequences of greater than 2000 bases provide us with the most accurate information concerning the phylogenetic relationships between sequence types, partial sequences dominate the entries that have been deposited into the international DNA databases. Using the alignment and analysis based on the "complete" *Rns* sequences we can evaluate partial sequences and place them into sequence types. We can also evaluate whether any unusual types exist for which no complete sequence has been obtained. By November 2013, there were 1487 partial *Rns* sequences in the DNA databases, almost 5 partial sequences for every complete sequence that had been determined. The increase each year in the partial sequences in the database is shown in Fig. 3.

Given the large number of both complete and partial sequences, we can examine the distribution of different sequence types in the databases. The distribution of sequence types in both complete and partial sequences is given in Table 1. It is clear that type T4 dominates the data, represented by 1300 sequences, representing more than 70% of all sequences in the databases. Only a single other sequence type, T5, even reaches 100+ representatives in the databases, but still constitutes only 5.7% of sequences. Four other sequence types (T3, T15 and T2–T6) exceed 2% of sequences in the collection. The combination of T2 and T6 needs to be considered in more detail, below. As mentioned, T15 (which is constituted of isolates classified as *Acanthamoeba jacobsi*) lacks even a single sequence that is viewed as essentially complete. The longest sequence (AY262365 – *A. jacobsi* AC305) is 1468 bases in length.

Examination of Table 1 in conjunction with the information represented in Fig. 2 provides considerable insight into the relationship between sequence types and their phylogenetic placement in the genus. We can begin by examining how sequence type and phylogenetic relationships relate to morphological groups in *Acanthamoeba*.

6. Morphological groups and sequence types: group I

The Group I *Acanthamoebae* represent the most differentiated members of the genus. With respect to sequence types, they are found towards the bottom of the tree, and constitute Types T7, T8, T9, T17 and T18, and in total represent slightly more than 2% of all *Acanthamoeba* sequences. Several species names have been applied to isolates that have been shown to have Group 1 morphology. Type T7 is represented by *Acanthamoeba astronyxis* (Ray & Hayes) (ATCC 30137). Type T8 includes *Acanthamoeba tubiashi* OC-15C (ATCC 30867). Type T9 appears heterogeneous, with the type strain being *Acanthamoeba comandoni* Comandon & de Fonbrune (ATCC 30135). However other T9 isolates have been described (incorrectly?) as *A. astronyxis* (ATCC 30901). T17 may be heterogeneous, with variation among isolates. The T17 sequence type currently has no isolates identified to a named species. Type T18 has only recently been reported, and has been given the new nominal *Acanthamoeba byersi* (Qvarnstrom et al., 2013). Given the relative rarity of Group I genotypes in the database, and their divergence from other *Acanthamoeba*, more work to characterize them is required. From our phylogenetic analyses, we believe that it is likely that this group deserves generic recognition, and should receive a generic tag such as *Megacanthamoeba*, given their larger trophozoite and cyst sizes compared to Group 2 and Group 3 *Acanthamoeba*.

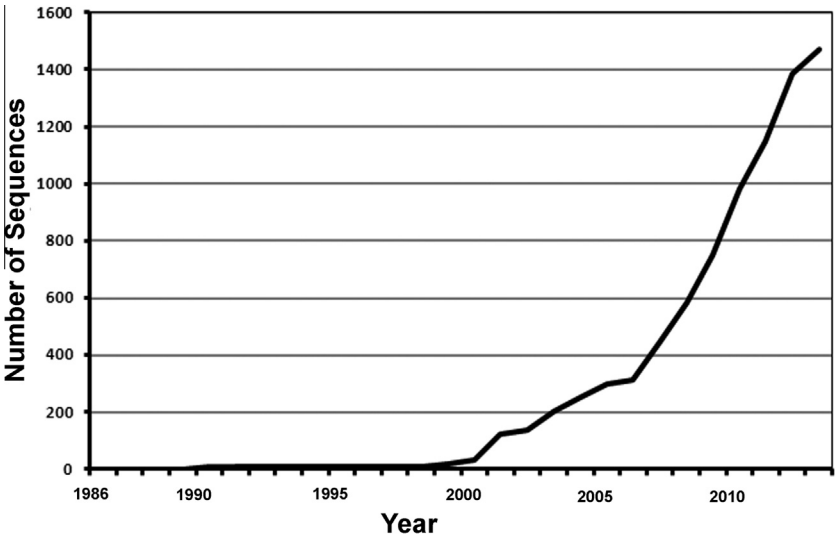


Fig. 3. Cumulative increase by year in the number of *Acanthamoeba Rns* sequences shorter than 2000 nucleotides in length in the international DNA databases through November 2013.

Table 1
Number of isolates and frequency of sequence types among data from the international DNA databases.

Sequence type [representative isolate (acc #)]	# Sequences >2000 nuc. (% of total)	# Sequences <2000 nuc. (% of total)
T1 – *A. castellanii* V006 (U07400)	2 – (0.6%)	12 – (0.8%)
T2* – *A. palestinensis* Reich (U07411)	13 – (3.9%)	28 – (1.9%)
T3 – *A. griffini* S7 (U07412)	7 – (2.1%)	86 – (5.8%)
T4* – *A. castellanii* ATCC 30011 (U07413)	235 – (70.6%)	1065 – (71.6%)
T5 – *A. lenticulata* Jc-1 (U94739)	14 – (4.2%)	90 – (6.1%)
T6* – *A. hatchetti* 11DS (AF251939)	10 – (3.0%)	61 – (4.1%)
T7 – *A. astronyxis* (Ray & Hayes) (AF019064)	3 – (0.9%)	4 – (0.3%)
T8 – *A. tubiashi* OC-15C (AF019065)	1 – (0.3%)	2 – (0.1%)
T9 – *A. comandoni* (AF019066)	6 – (1.8%)	4 – (0.2%)
T10 – *A. culbertsoni* Lilly A-1 (AF019067)	2 – (0.6%)	5 – (0.3%)
T11 – *A. hatchetti* BH-2 (AF019068)	7 – (2.1%)	24 – (1.6%)
T12 – *A. healyi* V013 (AF019070)	2 – (0.6%)	8 – (0.5%)
T13 – UWC9 (AF132134)	5 – (1.5%)	12 – (0.8%)
T14 – PN13 (AF333609)	2 – (0.6%)	0
T15 – *A. jacobsi* 31-B (AY262360)	0	41 – (2.8%)
T16* – U/HC1 (AY026245)	2 – (0.6%)	1 – (0.05%)
T17 – Ac E1a (GU808277)	5 – (1.5%)	8 – (0.5%)
T18 – CDC:V621 (KC822461)	10 – (0.3%)	0
T19 – OSU 04-020 (DQ451160)	1 – (0.3%)	16 – (1.0%)

* Sequence types T2, T4, T6 and T16 show evidence of significant subtypes.

7. Morphological groups and sequence types: group 3

Interpreting the information from phylogenetic analyses, and considering the species names that have been applied to various isolates, it appears that the ancestor of Group 1 *Acanthamoebae* and an amoebae that would evolve into a morphology equivalent to that defined for Group 3 diverged from the common ancestor of all members currently recognized as *Acanthamoeba*. In Fig. 2, Group 3 forms are represented by sequence types T2, T5, T6, T10, T12 and T14. I assume that sequence type T13 and T16 in Fig. 2 are also Group 2 forms, but I am unaware of morphological studies that confirm or refute this hypothesis. With this caveat, examination of Fig. 2 indicates that Group 3 is a paraphyletic grouping of sequence types with respect to the Group 2 *Acanthamoeba* represented by Types T3, T4 and T11. The isolates in sequence type T1 have been assigned species names placing them in Group 3, but that placement should be carefully evaluated. Type T19 is a new form just recently recognized, with a single complete sequence available for study, together with a number of partial sequences. It too should be evaluated for trophozoite and cyst form.

Types T13 and T16 are sister clades and very close to one another in sequence, but represent very rare types in the databases, and presumably in the environment. Type T13 is best known as forms that carry unusual bacterial endosymbionts (Horn et al., 1999).

Genotypes T2 and T6 are closely related, and much more common. Careful analysis indicates that members of the two types often do not meet the original 5% definition of separate genotypes in *Acanthamoeba*. Analysis of phylogenetic relationships indicates that the T2–T6 clade is heterogeneous, including five significant subtypes within a larger super-type Fig. 4. Sequences placed with Types T2 and T6 seem to represent a heterogeneous set of closely related sequences that actually constitute five sub-types within a larger super type. (This is somewhat similar to the more extensive case for Type T4 that will be discussed below.) The original T2 classification is associated with the Reich isolate *of Acanthamoeba palestinensis* (ATCC 30870), while the original T6 classification is associated with *A. palestinensis* 2802 (ATCC 50708). Relationship of species names with the other three subclades is unsettled, since all are only associated with "*Acanthamoeba polyphaga*" which is clearly a species name applied promiscuously throughout the genus. The final relationships between species names and genotype will have to be considered in a future paper. The three intermediate sub-types have been designated T2/6a, T2/6b and T2/6c (Fig. 4).

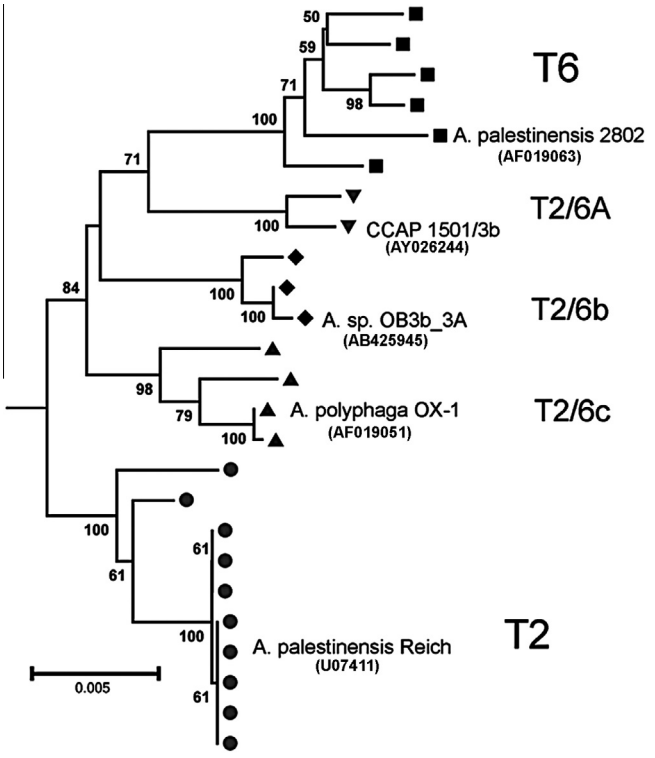


Fig. 4. Phylogenetic relationships among various representatives of sequence types T2 and T6 as determined from *Acanthamoeba Rns* sequences greater than 2000 bases. The tree was constructed using the neighbor joining method. Different symbols indicate members of particular sub-clades within the T2–T6 superclade. The sequences used as representative of each of the five sub-clades are labeled. Further analysis of the T2–T6 superclade appears in Fuerst et al. (2014).

8. Morphological groups and sequence types: group 2

Fig. 2 indicates that almost all of the species names traditionally associated with Group 2 *Acanthamoeba* fall into three Sequence types, T3, T4 and T11. This group appears to have diverged from within the Group 3 forms. Type T4 as mentioned previously is most common in the DNA databases, and this appears to reflect both its importance in the environment as well as for disease. It is certainly the most frequent sequence type associated with keratitis.

Table 2
Number of isolates and frequency of T4 sequence subtypes among data from the international DNA databases.

T4 subtype	Representative isolate	Accession #	Number of isolates in database
T4A	*A. castellanii* ATCC 30011	U07413	620
T4B	*A. castellanii* Ma ATCC 50370	U07414	252
T4C	Fernandez ATCC 50369	U07409	86
T4D	*A. rhysodes* Singh ATCC 30973	AY351644	172
T4E	*A. polyphaga* Page-23 ATCC 30871	AF019061	111
T4F	*A. triangularis* SH621 ATCC 50254	AF346662	22
T4-Neff	*A. castellanii* Neff ATCC 30010	U07416	34

Examination of the phylogenetic relationships between T4 sequences, however, suggests that this genotype is highly heterogeneous, and represents a series of moderately to well differentiated clades that do not meet the 5% criteria for Type status, but do represent monophyletic evolutionary lineages.

Subgroups have been assigned labels T4A through T4F and T4-Neff to represent seven sub-types within Type T4. These may well represent phylogenetically defined "species", especially since questions concerning sexual reproduction in *Acanthamoeba* leave application of the Biological species concept problematic. Table 2 lists a representative isolate with database accession number for each of the subtypes, as well as the number of isolates in the database that can be assigned to each subtype of T4.

As seen in Table 2, sub-type T4A is the most common form in the database. It is represented by the original isolate of *A. castellanii*. If one examines carefully each of the subtypes, it is not clear that any, with the possible exception of T4F, is monospecific. There are at least 14 species names that have been applied to isolates that are genotypically T4 and all of the subtypes contain multiple species names. Clearly, almost all of the species names that have been applied at one time or another within *Acanthamoeba* should be viewed as invalid until they are redefined.

A second major aspect of the subtypes of T4 concerns the isolate *A. castellanii* Neff. Even within the subtypes of T4, Neff represents a very small minority of isolates (34 of 1300). The numbers are actually smaller, since separate cultures of ATCC 30010 have been sequenced and deposited a number of times (perhaps as many as 5+ of the T4-Neff sequences are simply re-sequencing of ATCC 30010). The fact that this isolate has been chosen to represent the genus for the original genome sequence may be unfortunate, since genotypically it is not representative, and many of its characteristics are also not representative, even of T4.

9. Summary

Over 1800 submissions of full or partial sequences of the 18S rRNA gene of *Acanthamoeba* have been made to the international DNA databases. These represent only a part of the information that can be garnered for the analysis of the phylogenetic history of these amoebae. There are other genes, such as the mitochondrial 16-S like small subunit rRNA gene (Ledee et al., 2003) or the mitochondrial cytochrome-C oxidase subunit I gene (Crary, 2012) for which there is a much smaller, but growing set of sequences that can be compared to the data from the nuclear *Rns* gene. Only when we begin to incorporate multi-gene information will be completely satisfied with the answers to the question concerning "What is *Acanthamoeba*?", Future analyses will deal specifically with the questions of what species names should be assigned to sequence types and how data from different genes compare.

Acknowledgment

Thank you to all members of the *Acanthamoeba* community who have generously shared data that has been incorporated into this analysis. Further information is available at our website http://u.osu.edu/acanthamoeba.

References

Corsaro, D., Venditti, D., 2010. Phylogenetic evidence for a new genotype of *Acanthamoeba* (Amoebozoa, Acanthamoebida). Parasitol. Res. 107, 233–238.
Crary, M., 2012. Genetic Variability and its Relationship to *Acanthamoeba* Pathogenesis (Ph.D. thesis). The Ohio State University, Columbus, Ohio, pp. Xx.
Daggett, P.M., Sawyer, T.K., Nerad, T.A., 1982. Distribution and possible interrelationships of pathogenic and non-pathogenic *Acanthamoeba* from aquatic environments. Microb. Ecol. 8, 371–386.
Daggett, P.M., Lipscomb, D., Sawyer, T.K., Nerad, T.A., 1985. A molecular approach to the phylogeny of *Acanthamoeba*. Biosystems 18, 399–405.
De Jonckheere, J.F., 1983. Isoenzyme and total protein-analysis by agarose isoelectric-focusing, and taxonomy of the genus *Acanthamoeba*. J. Protozool. 30, 701–706.
Felsenstein, J., 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. 17, 368–376.
Fuerst, P.A., Booton, G.C., Crary, M., 2014. Phylogenetic analysis and the evolution of the 18S rRNA gene typing system of *Acanthamoeba*. J. Eukaryot. Microbiol. (in press).
Gast, R.J., 2001. Development of an *Acanthamoeba*-specific reverse dot-blot and the discovery of a new ribotype. J. Eukaryot. Microbiol. 48, 609–615.
Gast, R.J., Ledee, D.R., Fuerst, P.A., Byers, T.J., 1996. Subgenus systematics of *Acanthamoeba*: four nuclear 18S rDNA sequence types. J. Eukaryot. Microbiol. 43 (6), 498–504.
Gunderson, J.H., Sogin, M.L., 1986. Length variation in eukaryotic rRNAs: small subunit rRNAs from the protists *Acanthamoeba castellanii* and *Euglena gracilis*. Gene 44, 63–70.
Hewett, M.K., Robinson, B.S., Monis, P.T., Saint, C.P., 2003. Identification of a new *Acanthamoeba* 18S rRNA gene sequence type, corresponding to the species *Acanthamoeba jacobsi* Sawyer, Nerad and Visvesvara, 1992 (Lobosea: Acanthamoebidae). Acta Protozool. 42, 325–329.
Horn, M., Fritsche, T.R., Gautom, R.K., Schleifer, K.H., Wagner, M., 1999. Novel bacterial endosymbionts of *Acanthamoeba* spp. related to the *Paramecium caudatum* symbiont *Caedibacter caryophilus*. Environ. Microbiol. 1, 357–367.
Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.
Ledee, D.R., Booton, G.C., Awwad, M.H., Sharma, S., Aggarwal, R.K., Niszl, I.A., Markus, M.B., Fuerst, P.A., Byers, T.J., 2003. Advantages of using mitochondrial 16S rDNA sequences to classify clinical isolates of *Acanthamoeba*. Invest. Ophthalmol. Vis. Sci. 44, 1142–1149.
Lanocha, N., Kosik-Bogacka, D., Maciejewska, A., Sawczuk, M., Wilk, A., 2009. The occurrence *Acanthamoeba* (free living amoeba) in environmental and respiratory samples in Poland. Acta Protozool. 48, 271–279.
Nuprasert, W., Putaporntip, C., Pariyakanok, L., Jongwutiwes, S., 2010. Identification of a novel T17 genotype of *Acanthamoeba* from environmental isolates and T10 genotype causing keratitis in Thailand. J. Clin. Microbiol. 48, 4636–4640.
Pussard, M., Pons, R., 1977. Morphologie de la paroi kystique et taxonomie du genre *Acanthamoeba* (Protozoa, Amoebida). Protistologica 13, 557–598.
Qvarnstrom, Y., Nerad, T.A., Visvesvara, G.S., 2013. Characterization of a new pathogenic *Acanthamoeba* Species, A. byersi n. sp., isolated from a human with fatal amoebic encephalitis. J. Eukaryot. Microbiol. 60, 626–633.
Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.
Schroeder, J.M., Booton, G.C., Hay, J., Niszl, I.A., Seal, D.V., Markus, M.B., Fuerst, P.A., Byers, T.J., 2001. Use of subgenic 18S ribosomal DNA PCR and sequencing for genus and genotype identification of *Acanthamoebae* from humans with keratitis and from sewage sludge. J. Clin. Microbiol. 39, 1903–1911.
Stothard, D.R., Schroeder-Diedrich, J.M., Awwad, M.H., Gast, R.J., Ledee, D.R., Rodriguez-Zaragoza, S., Dean, C.L., Fuerst, P.A., Byers, T.J., 1998. The evolutionary history of the genus *Acanthamoeba* and the identification of eight new 18S rRNA gene sequence types. J. Eukaryot. Microbiol. 45, 45–54.
Visvesvara, G.S., 1991. Classification of *Acanthamoeba*. Rev. Infect. Dis. 13 (Suppl. 5), S369–S372.