Between a Pod and a Hard Test: The Deep Evolution of Amoebozoa

Seungho Kang,1,2 Alexander K. Tice,1,2 Frederick W. Spiegel,3 Jeffrey D. Silberman,3 Tomáš Pánek,4 Ivan Čepička,5 Martin Kostka,6,7 Anush Kosakyan,8 Daniel M.C. Alcântara,8 Andrew J. Roger,9 Lora L. Shadwick,3 Alexey Smirnov,10 Alexander Kudryavtsev,10 Daniel J.G. Lahr,8 and Matthew W. Brown*,1,2

1Department of Biological Sciences, Mississippi State University, Mississippi State, MS
2Institute for Genomics, Biocomputing & Biotechnology, Mississippi State University, Mississippi State, MS
3Department of Biological Sciences, University of Arkansas, Fayetteville, AR
4Department of Biology and Ecology, University of Ostrava, Ostrava, Czech Republic
5Department of Zoology, Charles University, Prague, Czech Republic
6Institute of Parasitology, Biology Centre of the Czech Academy of Sciences, České Budějovice, Czech Republic
7Department of Parasitology, University of South Bohemia, České Budějovice, Czech Republic
8Department of Zoology, University of São Paulo, São Paulo, Brazil
9Centre for Comparative Genomics and Evolutionary Bioinformatics, Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, Canada
10Department of Invertebrate Zoology, Faculty of Biology, Saint Petersburg State University, Saint Petersburg, Russia

*Corresponding author: E-mail: mwb250@msstate.edu.

Abstract

Amoebozoa is the eukaryotic supergroup sister to Obazoa, the lineage that contains the animals and Fungi, as well as their protistan relatives, and the breviate and apusomonad flagellates. Amoebozoa is extraordinarily diverse, encompassing important model organisms and significant pathogens. Although amoebozoans are integral to global nutrient cycles and present in nearly all environments, they remain vastly understudied. We present a robust phylogeny of Amoebozoa based on broad representative set of taxa in a phylogenomic framework (325 genes). By sampling 61 taxa using culture-based and single-cell transcriptomics, our analyses show two major clades of Amoebozoa, Discosea, and Tevosa. This phylogeny refutes previous studies in major respects. Our results support the hypothesis that the last common ancestor of Amoebozoa was sexual and flagellated, it also may have had the ability to disperse propagules from a sporocarp-type fruiting body. Overall, the main macroevolutionary patterns in Amoebozoa appear to result from the parallel losses of homologous characters of a multiphase life cycle that included flagella, sex, and sporocarps rather than independent acquisition of convergent features.

Key words: phylogenomics, transcriptomes, Amoebozoa, reductive evolution, phylotranscriptomics.

Introduction

Amoebozoa is the eukaryotic supergroup sister to Obazoa, the lineage that contains the animals and Fungi, as well as their protistan relatives, and the breviate and apusomonad flagellates. Amoebozoa is extraordinarily diverse, encompassing important model organisms and significant pathogens. Although amoebozoans are integral to global nutrient cycles and present in nearly all environments, they remain vastly understudied. We present a robust phylogeny of Amoebozoa based on broad representative set of taxa in a phylogenomic framework (325 genes). By sampling 61 taxa using culture-based and single-cell transcriptomics, our analyses show two major clades of Amoebozoa, Discosea, and Tevosa. This phylogeny refutes previous studies in major respects. Our results support the hypothesis that the last common ancestor of Amoebozoa was sexual and flagellated, it also may have had the ability to disperse propagules from a sporocarp-type fruiting body. Overall, the main macroevolutionary patterns in Amoebozoa appear to result from the parallel losses of homologous characters of a multiphase life cycle that included flagella, sex, and sporocarps rather than independent acquisition of convergent features.

Key words: phylogenomics, transcriptomes, Amoebozoa, reductive evolution, phylotranscriptomics.
FIG. 1. Representative trophic cells of amoebozoans examined in this study. (A) *Trichosphaerium* sp. leathery shell corycid amoeba. (B) *Diploclamysis* sp. leathery shell corycid amoeba. Scale bar = 50 μm. (C) *Amphizonella* sp. leathery shell corycid amoeba. Scale bar = 50 μm. (D) *Micriamoeba* sp. amoeba. (E) *Echinamoeba exudans* amoeba. (F) *Vermamoeba vermiformis* amoeba. (G) *Vermamoeba* sp. (CCAP1503-5) amoeba. (H) *Rhizamoeba saxonica* amoeba. (I) *Flabellula citata* amoeba. (J) *Cryptodifflugia operculata* testate amoeba. (K) *Diffugia bryophila* testate amoeba. Scale bar = 50 μm. (L) *Arcella intermedia* testate amoeba. (M) *Nolandella* sp. amoeba. (N) *Copromyxa protea* amoeba. (O) *Amoeba proteus* amoeba. Scale bar = 100 μm. (P) *Squaamamoeba japonica* amoeba. Scale bar = 5 μm. (Q) *Rhizomastix elongata* amoeboflagellate. (R) *Pelomyxa* sp. amoeboflagellate. Scale bar = 50 μm. (S) *Mastigella elihardi* amoeboflagellate. (T) *Mastigamoeba abducta* amoeboflagellate. (U) *Echinosteliospis oligospora* amoeba. (V) *Echinostelium minutum* amoeboflagellate. (W) *Echinostelium bisporum* amoeboflagellate. (X) *Protosporangium articulatum* amoeboflagellate. (Y) *Clastostelium recurvatum* amoeba. (Z) *Flamella aegyptia* amoeba. (AA) *Phalansterium* Deep Evolution of Amoebae. doi:10.1093/molbev/msx162
in supplementary table S1, Supplementary Material online), sexual states (Brown et al. 2007; Lahr et al. 2011b; Spiegel 2011; Tekle et al. 2017), and/or spore-bearing structures called fruiting bodies (fig. 2; Shadwick et al. 2009; 2016; Kudryavtsev et al. 2014).

Fruiting bodies among amoebozoans produce dormant, walled propagules (spores, see further definition in supplementary table S1, Supplementary Material online) that are derived through the development of these structures. Sorocarps like those in the well-studied Dictyostelium are formed by aggregation of individual cells that work in concert to form an emergent, aerial multicellular structure (fig. 2A) with many spores (Brown and Silberman 2013). This trait evolved at least seven times independently in eukaryotes (Brown and Silberman 2013). In contrast, sporocarps are formed by a single cell that produces an elevated extracellular, subaerial stalk upon which it develops into a spore or cleaves into several spores (fig. 2B–S; Olive 1975; Shadwick et al. 2009; Tice et al. 2016a). Unlike sorocarps, sporocarpic development is unique to Amoebozoa, where it is found in the microscopic protosteloid amoebae as well as in the mostly macroscopic myxogastrial plasmodial (i.e., large multinucleate cell) slime molds (see Olive 1975; Shadwick et al. 2009).

It is has not been well established if the various types of amoeboid cells among Amoebozoa are homologs of each other (Spiegel and Feldman 1985; Spiegel et al. 1995). In addition to being trophozoites (feeding cells), some amoeboid cells may function as gametes or zygotes or they may become committed to differentiate into fruiting bodies or resting stages (Spiegel and Feldman 1985).

Amoebozoans evolved from their last common ancestor around 1.2 billion years ago (Erne et al. 2014), and as such their evolutionary relationships are deep and are currently not well understood. Without such understanding, it is difficult to discern the macroevolutionary patterns responsible for the extreme diversity in cell types and life cycles within this supergroup. The morphological characters, alone, are considered too few and ambiguous to accurately resolve a phylogeny over the evolutionary distances in Amoebozoa (Page 1988; Patterson 1999). However, early molecular phylogenetic analyses, based mostly on nuclear small subunit ribosomal RNA gene sequences (SSU), yielded no better insights than morphology because they suffered from very poor taxon sampling and more issues (i.e., extreme rate differences and
compositional bias). With increasing taxon sampling and/or number of genes analyzed, the monophyly of what we now call Amoebozoa began to emerge (Bolivar et al. 2001; Fahni et al. 2003; Cavalier-Smith 1998; Pawlowski 2008; Shadwick et al. 2009; Lahr et al. 2011a), and many, robust lower-level taxa that mostly correspond with classical orders have been established (Amaral Zettler et al. 2000; Tekle et al. 2008; Lahr et al. 2011a; Smirnov et al. 2011). These early analyses hinted at the existence of several higher, supra-ordinal taxa (Cavalier-Smith et al. 2004; Smirnov et al. 2005), although only one of these (Tubulinea) appears to be well supported on a consistent basis (Cavalier-Smith et al. 2004; Lahr et al. 2011a; Smirnov et al. 2011; Berney et al. 2015). The remaining deepest branches in the Tree of Amoebozoa have remained very difficult to resolve (Cavalier-Smith et al. 2016; Tekle et al. 2016).

Recent phylogenomic studies strongly support the monophyly of Amoebozoa (Brown et al. 2013), but robust inferences on the deepest relationships and even the composition of proposed major lineages are likely compromised by a severe undersampling of the known amoebozoan diversity (Cavalier-Smith et al. 2015, 2016; Tekle et al. 2016). Taxa underrepresented in these analyses include those that have 1) complex life cycles involving the sequential development of multiple amoeboid morphologies, 2) flagellated cells, 3) sporocarps, and 4) tests. Here, we present data derived from a set of 325 protein-coding genes drawn from a sampling of 86 amoebozoans that represent the known morphological diversity of the supergroup in its entirety (figs. 1, 2). We include taxa from previous phylogenomic studies that focused on Amoebozoa (Cavalier-Smith et al. 2015, 2016; Tekle et al. 2016; Tice et al. 2016a), and collected new transcriptomic data for 61 additional species (figs. 1–3, and robust site sampling per taxon (fig. 3, histogram, see supplementary table S2, Supplementary Material online). Our transcriptomic data include the archetypal amoebozoan, A. proteus, nearly all genera for which a flagellate state is known, most taxa that are known or suspected to be sexual (Lahr et al. 2011b), several sporocarpic myxogastroids, and almost all genera that are known to produce unicellular stalked, protosteloid fruiting bodies via sporocarp (Spiegel 1990; Shadwick et al. 2009). We also include representatives of both clades where sporocarpic fruiting is known (Brown et al. 2011; Romeralo et al. 2013) and both known types of testate amoebzoans, the leathery-shelled amoebae (herein named corycid amoebae) and the rigid-shelled arcellinids. To a great extent, our capacity to sample such taxonomic depth and diversity is due to our utilization of single cell transcriptomic methods. This allowed us to include species in our analyses that have not been cultivated or are predators of other eukaryotes (see supplementary table S3, Supplementary Material online).

We robustly recover for the first time major lineages of amoebzoans (Evosea, Tubulinea, and Discosea) and a highly supported dichotomy between Tubulinea + Evosea (Tevosa) and Discosea. We can now infer that the last common ancestor of the group most likely had a multistate life cycle with sex, flagella, and probably sporocarps. From these data and the mapping of these phenotypic traits, we are able to develop several testable hypotheses concerning the macro-evolutionary patterns of morphology in Amoebozoa, which will enable subsequent studies to test our hypotheses.

Results

A Resolved Tree of Amoebozoa

For our phylogenomic results presented herein, a supermatrix composed of 325 proteins from 98 taxa was constructed. Initial analyses of this full supermatrix with simple site-homogeneous (LG+Γ4+F; see supplementary fig. S1, Supplementary Material online) and more sophisticated site-heterogeneous (LG+Γ4+F+C60+PMSF; see supplementary fig. S2, Supplementary Material online) models (detailed below) conflicted regarding several deep nodes of the Amoebozoa tree. For example, the Tevosa clade was recovered by the site-heterogeneous model (see supplementary fig. S2, Supplementary Material online), but not by the site-homogeneous model (see supplementary fig. S1, Supplementary Material online).

To examine this conflict, we estimated the rate of evolution at sites in the supermatrix and progressively removed them in a stepwise fashion, plotting bootstrap values for nodes of interest in the tree per site deletion step under the models LG+Γ4+F and LG+Γ4+F+C60+PMSF models estimated in IQ-Tree (fig. 4A, 8). The rationale for this procedure is that the fastest evolving sites within a deep phylogenetic analysis are often saturated with multiple substitutions and, as a result of model-misspecification can manifest nonphylogenetic signal especially when overly simplistic models are used (Jeffroy et al. 2006; Olsen 1987; Lartillot and Philippe 2008; Brown et al. 2013). Analysis of the numbers of substitutions per sites in the various deletion data sets revealed that the full data set displayed substitutional saturation and removal of fastest evolving sites did ameliorate the problem (fig. 4C). With the progressive removal of the fastest sites, the site homogeneous analyses yielded increasing support for the same major groups as the initial site-heterogeneous analyses. In contrast, the topology estimated by site-heterogeneous model and its support values were unaffected by fast-site removal until well over half the data set was removed (fig. 4B). To avoid problems associated with model misspecification in fast-evolving sites, all of our subsequent phylogenomic analyses were based on a data set that excludes fastest evolving sites as well as the removal of uninformative constant sites (fig. 4, the 17,500 fast site removal data set, less constant sites). This noise reduction step also had the benefit of reducing the data set size, lessening the extreme computational burden of phylogenomic tree inference under Phylobayes, which is a particularly important consideration when working with data sets of this size. For example, our Phylobayes analyses presented herein required ~979,200 CPU hours.

The resulting Tree of Amoebozoa, in figure 3 is based on the Bayesian site-heterogeneous mixture model, CAT-GTR, the most realistic phylogenetic model available (Lartillot et al. 2013). Although the CAT-GTR model cannot be used in the ML framework, the LG+Γ4+F+C60+PMSF model that similarly estimates site-specific amino acid (AA)
profiles for phylogenomic analyses based on the C60 empirical frequency profiles (Le and Gascuel 2008) was used for ML analyses in IQ-TREE v1.6j under the CAT + GTR model of protein evolution, with two converged independent chains, burnin 1,850 generations, with a postburnin sampling of ~2,800 generations. Values at nodes are posterior probability and ML bootstrap (MLBS) (1,000 ultrafast BS reps, IQ-TREE 94% 100% ML Bootstrap = 1.0 Bayesian PP = 1.0 Bayesian PP). Two converged PHYLOBAYES-MPI methods.

As with many previous molecular phylogenetic analyses (Cavalier-Smith 1998; Pawlowski 2008; Shadwick et al. 2009; Lahr et al. 2011a), we find that Amoebidae is a fully supported clade. Armed with strong statistical support for the deepest nodes within the Tree of Amoebidae we show that there are three major lineages. The lineages Tubulina and Discosea continue to be recovered as previously suggested (Cavalier-Smith et al. 2015, 2016; Tice et al. 2016a). However, we do not recover Lobosa sensu Smirnov et al. (2011), that is,
a clade comprising Discosea + Tubulinea as sister lineages or
Lobosa sensu Cavalier-Smith et al. (2016) (Discosea + Tubulinea + Cutosea, listed as Lobosa in fig. 4 and table 1). Additionally, constraining the tree with Lobosa (Discosea + Tubulinea, with or without Cutosea) can be rejected under approximately unbiased (AU) tests (Shimodaira, 2002) at a confidence interval of 95% using our data set (P-value = 0.0059 and 0.0014, respectively, table 1). Instead Tubulinea is sister to a major monophyletic lineage we call Evosea (named herein composed of Eumycetozoa, Variosea, Archaeamoebae, and Cutosea) (BPP = 1.0, MLBS = 99%, fig. 3, see supplementary fig. S3, Supplementary Material online). We propose the name Tevosa for the clade Tubulinea + Evosea.

These backbone clades of Amoebozoa are in concordance with Tekle et al. (2016), but not with Cavalier-Smith et al. (2016). The deepest clades we confidently accept in Amoebozoa are Tevosa and Discosea. The group names we
are using and have named herein are mapped on the tree in figure 3 and are listed with a novel clade in the supplementary text, Supplementary Material online. We have modified the composition and definitions of all major lineages with respect to our findings. On the basis of the phylogenomic evidence, the deepest node in Amoebozoa appears to be between Discosea and Tevosa. Our results lead us to reject the concepts of Lobosa sensu Smirnov et al. (2011) (Discosea + Tubulinea) and Lobosa sensu Cavalier-Smith et al. (2016) (Discosea + Tubulinea + Cutosea) (fig. 3, table 1). Because the clade Tevosa has no obvious unifying morphological traits, we will discuss Tubulinea and Eovosea separately.

**Tubulinea**

Most members of Tubulinea have an amoeboïd state in their life histories that is tubular in cross section and lacks pseudopodia, though a few have a flat cross section and subpseudopodia (Smirnov et al. 2011; fig. 1A–O, see supplementary table S1, Supplementary Material online). We recover Tubulinea, a group that often shows strong statistical support even when few genes are analyzed (Lahr et al. 2011a, 2013; Cavalier-Smith et al. 2015, 2016). The subclades of Tubulinea are all strongly supported by both BPP and MLBS. With our broad sampling of Tubulinea taxa, we provide deep resolution and reveal a novel subclade, which we name Corycidia (fig. 3, see supplementary text, Supplementary Material online). Corycidia is the strongly supported sister lineage to the rest of the Tubulinea (includes Echinamoebidia and Elardia [see supplementary taxonomic summary, Supplementary Material online]). Corycidian taxa are characterized by having a leathery flexible tests (fig. 1A–C), distinct from the rigid tests found within Arcellinida (fig. 1J–L, see supplementary table S1, Supplementary Material online). Corycidia includes Diplochlamys sp. (fig. 1B), Amphizonella sp. (fig. 1C), and *Trichosphaerium* sp. (fig. 1A), which were difficult to place within Amoebozoa in previous phylogenomic studies due to poor taxon sampling, specifically from within Tubulinea (Tekle et al. 2008; Cavalier-Smith et al. 2016; Tekle et al. 2016). In previous, less comprehensive studies, the only member of Corycidia included was *Trichosphaerium* sp.; it was placed either within Tubulinea (Cavalier-Smith et al. 2016) or sister to Tubulinea (Tekle et al. 2016). This particular isolate of *Trichosphaerium* was renamed “*Atrichosa algivora*” in (Cavalier-Smith et al. 2016); however, given our observations of this isolate (fig. 1A) it likely represents a spicule-less stage (gamont) of *Trichosphaerium* the details of which are discussed in (Page 1983). A spicule-bearing *Trichosphaerium* must be examined further to confirm this.

**Eovosea**

The clade we name Eovosea contains the well-supported subclades Cutosea, Archamoebae, Eumycetozoa, and Variosea (fig. 3). However, we are hesitant to state categorically the exact branching orders among them because of the lack of MLBS support for the deepest nodes within the group (fig. 3, see supplementary fig. S3, Supplementary Material online) and AU test results show no significant resolution of the branching order (table 1). We demonstrate that Cutosea (Cavalier-Smith et al. 2016) belongs within this group and not within Lobosa (i.e., Tubulinea + Discosea) as argued in (Cavalier-Smith et al. 2016), in part because the proposed group Lobosa is not recovered in our analyses. Eovosea corresponds, more or less, with the traditional Conosa (Cavalier-Smith 1998) plus the Cutosea. However, Conosa is morphologically defined as having taxa with flagella associated with a radiating cone of microtubules emerging from either their anteriorly directed basal body or a microtubule organizing center associated with this basal body. As far as currently known, Cutosea is devoid of flagellated taxa (Cavalier-Smith et al. 2016). We feel that the most robust result and conclusion should be to provide a new name for this clade and not to subsume Cutosea into Conosa. We feel that leaving Conosa as a valid and nonsynonymous clade to Eovosea is a more reliable and stable taxonomic option. More work should be focused on this clade, particularly, isolation and transcriptomic sequencing efforts to collect more data from cutosean taxa.

Cutosea is represented by *Sapocribrum*, *Squamamoeba*, and American Type Culture Collection (ATCC) strain PRA-29 (deposited as “*Pessonella*”). The former two taxa are small
amoebae covered with very small scales (Pussard 1973; Tekle et al. 2008; Kudryavtsev and Pawlowski 2013; Lahr et al. 2015). ATCC PRA-29 was misidentified as Pessonella sensu (Pussard 1973; Tekle et al. 2008), and further work should be done to examine this strain in order to formally describe it.

Cutosea sensu Cavalier-Smith et al. (2016) was originally placed as sister to the rest of Lobosa. However, since their Amoebobooza-only trees (Cavalier-Smith et al. 2016) are unrooted, a topology where Cutosea is sister to Conosa (Archamoebae, Eumycetozoa, and Variosea) cannot be ruled out. In Tekle et al. (2016) the authors treat Cutosea as sister to Himatismenida, which, together, are sister to Tubulinina in their analyses. This result of Tekle et al. (2016) is probably due to a long branch attraction artifact that was not remedied by a more realistic evolutionary model such as the site heterogeneous model used in our analyses. Here, Cutosea appears to be sister to the rest of Evosea, that is, the traditional Conosa (fig. 3). In the full data set, there is full MLBS support for Conosa, but the order among Archamoebae, Eumycetozoa, and Variosea, with or without Cutosea branching within, is ambiguous when fast evolving sites are removed under the LG + F 4 + F + C60 + PMSF model (fig. 48).

We rename the well-supported group that contains myxogastrids, dicytostelids, and protosporangids (represented here by Ceratiomyxa, Clastostelum, and Protosporangium) Eumycetozoa rather than the recently coined term Macromycetozoa (Fiore-Donno et al. 2010) for two major reasons. First, Eumycetozoa is the older name such that it even has priority over the entirety of Amoebobooza (see Shadwick et al. 2009; Adl et al. 2012). However, the conservation of Amoebobooza was argued for because of its literature familiarity in Adl et al. (2012). Secondly, Eumycetozoa is a group that should include the myxogastrids and our usage corresponds to the Eumycetozoa hypothesis (Olive 1975) that posits a monophyletic group of exclusively fruiting protists that includes myxogastrids, dicytostelids, and some protosteloid amoebae, in this case, the protosporangids (fig. 3).

The previous incerta sedis protosteloid amoeba, Echinosteliospis oligospora (Spiegel 1990), a species that lacks a flagellate state in its life cycle and has an unusual multinucleolate nucleus (Reinhardt 1968) is a myxogastid slime mold (fig. 3). As previously suggested (Whitney et al. 1982), we confirm that the protosteloid amoeba Echinosteliospis bispornum (figs. 1W, 2C), which lacks a plasmodial state (Spiegel and Feldman 1989), is also a myxogastid and is sister to the more typical Echinostelium minutum (figs. 1V, 2B).

Our recovery of Variosea conforms with the hypothesis of Berney et al. (2015), although they were unable to show deep resolution with their 18S rRNA gene-based trees. Notably, we still lack the key multiflagellate variosean, Multicilia (Nikolaev et al. 2006); however, we conclusively demonstrate the monophyly of Variosea. Although some ambiguities exist between ML and Bayesian results (fig. 3, see supplementary fig. S1, Supplementary Material online) the backbone topology of Archamoebae corresponds well with the hypotheses about the relationships within this exclusively anaerobic clade, with a dichotomy yielding Pelobionta and Entamoebida (Pánek et al. 2016).

Evosea includes amoebae that are either tubular or flat in cross section (see supplementary table S1, Supplementary Material online). Archamoebae, Eumycetozoa, and Variosea each contain some species that are flagellated (usually amoeboflagellates sensu Spiegel 1990), and at least some members of all three groups (Adl et al. 2012) have flagella that contain an electron-dense element in their transition zone, a character not found in any other eukaryotes (see Spiegel 1990, 1991; Walker et al. 2001). Sex has been demonstrated in Eumycetozoa (in both Dictyostelia and Myxogastria) and life cycles consistent with sex (plasmogamy or uninucleate cells and two divisions similar to meiosis and/or obligate amoebae sensu Spiegel and Feldman (1985) arising from amoeboflagellates or genetic evidence of possible recombination) are present in several varioseans (e.g., Cavostelium aphophysatum, Ceratiomyxa tahitiensis; Spiegel 1990) as well as in Archamoebae (i.e., Entamoeba histolytica; Lahr et al. 2011b).

In protosteloid evoseans, sex is sometimes thought to be associated with life cycles that include amoeboflagellates that germinate from spores and precede the development of nonflagellate obligate amoebae that subsequently produce the sporocarps (Spiegel and Feldman 1985; Spiegel et al. 1995; Spiegel 2011; see supplementary table S1, Supplementary Material online). However, additional work on sex in Evosea should be conducted to examine the true nature of these life cycles.

Discosea

Discosean amoebzoaons are relatively flat in cross section, though some can be somewhat dome shaped, and they may or may not exhibit subpseudopodia (Brown et al. 2007; Smithm et al. 2013; Shadwick et al. 2016; Tice et al. 2016). Many of the flabellinids are fan-shaped and mostly uniaxial, though some can be multiaxial. Protosteloid, sporocarpic fruiting is found in both Flabellinia and Centramoebida (fig. 3 and see supplementary table S1, Supplementary Material online). Sex is suspected in the flabellinid Sappinia (Brown et al. 2007; fig. 3 and see supplementary table S1, Supplementary Material online). No sporocarp or flagellated taxa are found in Discosea (fig. 3 and see supplementary table S1, Supplementary Material online).

Discosea is fully supported in our analyses and is divided into two fully supported groups, Flabellinia and Centramoebida sensu Tekle et al. (2016). Even with their paucity of taxon sampling, the analyses of Cavalier-Smith et al. (2016) correspond with this dichotomy. In Cavalier-Smith et al. (2016) they recover Centramoebida and Himatismenida in their Centramoebia, but they did not sample any taxa in Pellitida (here represented by the genera Pellita, Gocevia, and Endostelium). All the species in their Flabellinia that correspond with species in our Flabellinia occur within the taxon, but the poorly supported relationships they inferred within the group are quite different than those we recover. Although the analyses of (Tekle et al. 2016) focused on Discosea, they did not recover the group as monophyletic. Their interpretation of results was plagued by incorrectly identified taxa, “Mayorella sp.” (strain BSH, MMETSP0417), which is a Cunea sp. (Cavalier-Smith et al. 2015), and “Pessonella sp.” (strain PRA-29, MMETSP0420), which is not true Pessonella. An assemblage
that they called Eudiscosea contains the Centramoebida (including one pellicid, Gocevia fonbrunei) as the sister to a clade that corresponds very closely to our interpretation of Flabellinia. However, they placed the Himatismenida as sister to the Cutosea, which together group as sister to a severely undersampled Tubulinea.

**Discussion**

A well-resolved rooted tree allows us to construct testable hypotheses concerning the macroevolutionary patterns of morphology that characterize Amoebozoa. Here, we have mapped life cycle characters onto the Tree of Amoebozoa (fig. 3) and list these characters by species in supplementary table S1, Supplementary Material online. On the basis of parsimony principles we can assess the macroevolutionary trends across the Tree of Amoebozoa, as well as the nature of the last common ancestor of Amoebozoa (LCAA; fig. 3, table S1). For instance, it is possible to conclude that LCAA had a flagellate state in its life history (Spiegel 1991, 2011; Spiegel et al. 1995; Lahr et al. 2011b; Adl et al. 2012; Yubuki and Leander 2013) as this character must have been present in the Last Eukaryote Common Ancestor (LECA; Goodenough and Heitman 2014). Relatively few amoebozoan lineages have a flagellate state, and all confirmed flagellate taxa are found in Evesoa (fig. 3, see supplementary table S1, Supplementary Material online; Spiegel 1991; Spiegel et al. 1995; Mikrjukov and Mylnikov 1998; Smirnov et al. 2011; Adl et al. 2012; Patakova et al. 2013; Berney et al. 2015; Zadrobilkova et al. 2015; Pánek et al. 2016).

Although some exceptions exist (e.g., Multicilia [Nikolaev et al. 2006] and some species of Phalansterium [Smirnov et al. 2011]), almost all amoebozoans have an amoeboid state in their life history. It is likely LCAA had an amoeboid state, but the evolution of particular types of amoebae could have been quite complex (Spiegel et al. 1995). Alternatively, if the ancestor had a complex life cycle with one amoeboid state that alternated with another, as in myxogastrids and several protosteloid amoebae (Spiegel and Feldman 1985), an interesting hypothesis presents itself. For example in Amoebozoa, amoebae can be divided into two major morphological categories, tubular in cross section with axial cytoplasmic flow, and flattened in cross section with more irregular cytoplasmic flow (Smirnov et al. 2005, 2011; see supplementary table S1, Supplementary Material online). Several amoebozoans have life cycles with amoeboid cells that can assume both morphologies at alternate stages (see supplementary table S1, Supplementary Material online), for example, the myxogastrids and some protosteloid amoebae (Olive 1975; Adl et al. 2012). Other amoebozoans, such as some tubulinean taxa (e.g., the Leptomixida—Rhizamoeba, Leptomixyx, and Flabellula; Page 1988), can transition back and forth in the same cell. It is clear from our results that the myxogastrids contain examples of the loss of the tubular plasmodial state seen in E. bisporum (Spiegel and Feldman 1989) and the loss of the alternate flagellate states found in E. oligospora (Reinhardt 1968).

Outside of Evesoa, Trichosphaerium (here shown to be in Corycidia in Tubulinea), reportedly has a flagellate state (Schmarda 1871), which should be further investigated. Nonetheless, most flagellated amoebozoans, barring a few derived taxa (e.g., Pelomyxa [Seravin and Goodkov 1987]), have typical eukaryotic axonemes (a 9 × 2 + 2 microtubule configuration) and basal bodies (a 9 × 3 microtubule configuration). These nearly universal features of flagella are consistent with their shared evolutionary history with LECA; thus, the presence of a flagellate state in LCAA. Flagellates in Archamoebae have a less complex flagellar apparatus compared with those within Variosea and Eumycetozoa (Patakova et al. 2013; Zadrobilkova et al. 2015; Pánek et al. 2016), and careful work has yet to conclude what homologies exist between the flagellar rootlets in Archamoebae and those of the more complex flagellar apparatuses in Variosea and Eumycetozoa. Some superficial comparisons have been made (Cavalier-Smith 1998; Cavalier-Smith et al. 2004, 2015, 2016). Since most rootlet elements have apparent homologs to those outside Amoebozoa (Spiegel 1991; Yubuki and Leander 2013), none is considered synapomorphs of the group, though their overall conformation might be synapomorphic. However, one potential synapomorphy of Amoebozoa, or at least in examined Evesoa, is the presence of an electron dense plug in the flagellar transition zone of many Eumycetozoa, Variosea, and Archamoebae (Spiegel 1990, 1991; Patakova et al. 2013). The presence of this character outside of Evesoa can only be evaluated if flagellates of Trichosphaerium are rediscovered and examined.

There are members within each of the deep amoebozoan clades that have life cycles indicative of sex (fig. 3, see supplementary table S1, Supplementary Material online; Lahr et al. 2011b) and sex is relatively common outside of Amoebozoa. Therefore, LCAA must have been sexual (Spiegel 2011; Tekle et al. 2017). Since sex does not appear necessary for reproduction in any amoebozoan, it was likely facultative in this ancestor. It is interesting to note that the stages of sex, or suspected sex, are often associated with transitions from one somatic state to another, for example, the alternation between amoebodynamic and plasmodial, a form of obligate amoeba, in the myxogastral slime molds and between amoebodynamic and obligate amoebae in protosteloid varioseans and eumycetozoa (Martin and Alexopoulos 1969; Spiegel and Feldman 1985; Spiegel et al. 1995).

Most amoebozoans, but not all (e.g., A. proteus [Page 1988]), make resting cysts that are different from spores in that they are sessile (i.e., not elevated above the substratum via a stalk as in protosteloid and myxogastral amoebae), walled, dormant cells (see supplementary table S1, Supplementary Material online). This is also true of most protistan groups. Moreover, most sporocarpic amoebae are also capable of producing sessile cysts that are morphologically distinct from spores (Tice et al. 2016a). We however do not yet know the genetic basis for encystment across the breadth of Amoebozoa, which leaves us ill-informed about whether all cysts are homologous throughout the group as well as to outgroup protistan taxa. Future developmental studies should be undertaken to examine this more fully. Nevertheless, we hypothesize that LCAA was capable of encystment.
Deep Evolution of Amoebae

Sporocarpy, where a single cell develops into a subaerial, stalked fruiting structure that bears spores, is unique to Amoebozoa (Shadwick et al. 2009). Other than amoebae and cysts, both of whose genetic basis need to be worked out to confirm if the various types of amoebae and the types of cysts are homologs throughout the group, the most widespread developmental state found across Amoebozoa is sporocarpy (figs. 2, 3, see supplementary table S1, Supplementary Material online; Olive 1975; Spiegel 1990; Shadwick et al. 2009; Kudryavtsev et al. 2014; Berney et al. 2015; Tice et al. 2016a). The most evolutionarily common type of sporocarpy across the Tree of Amoebozoa is protosteloid sporocarpy, where the cell that develops into a sporocarp and the sporocarp itself are microscopic and contain only one to a few spores (fig. 2B,D–5; Shadwick et al. 2009). Variosea and Eumycetozoa contain members that have sporocarpic fruiting (protosteloid and myxogastrid; fig. 2B–5, see supplementary table S1, Supplementary Material online). Protosteloid amoebae are present in both Evosea (Variosea and Eumycetozoa) and Discosea (Centramoebia and Flabellinia; fig. 3). If the root of Amoebozoa is between Discosea and the rest of the amoebozoans as suggested in our results, we may hypothesize that the last common ancestor of the whole group was sporocarpic, a possibility previously suggested (Shadwick et al. 2009). This hypothesis was recently rejected in Cavalier-Smith et al. (2016), where the authors suggested that protosteloid sporocarpy is simply the result of the addition of stalks to already existing cyst stages and could easily have evolved independently several times. However, in most protosteloid amoebae, cysts, when present, are morphologically distinct from spores (Olive 1975; Spiegel and Feldman 1993; Tice et al. 2016a). Also, fruiting body ultrastructure is remarkably similar in protosteloid amoebae in both Evosea and Discosea (Spiegel et al. 1979; Olive et al. 1984; Spiegel and Feldman 1993). Our hypothesis of the origin of sporocarpy is quite testable. Future work using comparative developmental transcriptomics will help us to determine if the molecular mechanisms contributing to sporocarpy have a common evolutionary basis. Should sporocarpy in Evosea and Discosea prove homologous, that would support the hypothesis that it was present in their common ancestor, and thus, given our phylogeny, LCAA.

Our data clearly demonstrate that it is reasonable to assume that most macroevolutionary patterns in Amoebozoa are likely the result of the loss of characters that were present in a complex last common ancestor that was sexual and flagellate and perhaps had more than one somatic stage and the ability to disperse propagules using sporocarpy. Nonetheless, there has also been the evolution of novel traits within the group: once in Copromyxa (Brown et al. 2011; fig. 2A, see supplementary table S1, Supplementary Material online) in Elardia (Tubulinea) and also in the developmentally distinct dictyostelids in Eumycetozoa (Romeralo et al. 2013; Evosea; fig. 3; see supplementary table S1, Supplementary Material online). It is interesting to note that both tests and sorocarpy are not unique to Amoebozoa and have convergently evolved in other major lineages of amoeboid eukaryotes. Sorocarpy is known in Heterolobosea (Acrasis), Opisthokonta (Foncicula), Rhizaria (Guttulinopsis), Stramenopiles (Sorodiplophrys; Brown and Silberman 2013, Tice et al. 2016b). Testate amoebae are also present in Stramenopiles and widely distributed in Rhizaria (Pawlowski 2008).

With this well-supported phylogenetic hypothesis based on many genes and covering the taxonomic and developmental breadth of Amoebozoa, it has finally become possible to begin to address the macroevolutionary patterns of amoebozoans in an objective and non speculative manner. Robust studies of this nature provide a starting point for evolutionary developmental approaches to systematically examine the homology of the proposed characters. Additionally, with the data presented here, we can begin to identify convergent and homologous characters and character-states, not only morphological but also genomic, that have shaped and consequently led to the evolution of this vastly diverse and ecologically important supergroup of eukaryotes.

Materials and Methods

Details of experimental methods for isolation, identification, culturing, microscopic methods, nucleic acid extraction, cDNA construction, Illumina sequencing, cluster assembly, phylogenomic matrix assembly, and fast evolving site removal from the phylogenomic matrix were performed as in Tice et al. (2016a) and are described in the supplementary text 1, Supplementary Material online. Methods associated with our AU tests are also described in the supplementary text 1, Supplementary Material online.

Phylogenomic Tree Inference

Concatenated matrices compiled of 325 genes from 98 taxa resulted in an alignment of 93,478 amino acid (AA) sites. From this data set a subset of the fastest evolving sites (17,500 AA sites; fig. 3, discussed below and in supplementary text, Supplementary Material online) and constant sites (12,821 AA sites), which do not contribute to phylogenetic signal, were removed, resulting in our primary matrix of 63,157 AA sites. Bayesian inferences were performed in PhyloBayes-MPI v1.6j (Lartillot et al. 2013). To account for site heterogeneous amino acid substitutions, we used CAT-GTR (Lartillot et al. 2013). For Bayesian analyses we ran four independent Markov chain Monte Carlo chains for ~4,500 generations. Two of the chains converged at 1,850 generations with the largest discrepancy in posterior probabilities (PPs) (maxdiff) <0.069. The consensus of the two converged chains is presented in figure 3. While the other three chains did not converge, their overall topology was largely congruent with our converged chains and a consensus tree of all four chains is presented in supplementary figure S4,
Supplementary Material online. ML trees were inferred in IQ-TREE v. 1.5.0 (Nguyen et al. 2014). IQ-TREE is currently the only high-performance ML program capable of implementing C-series models, which offer more realistic phylogenetic site-heterogeneous models, that are not options in other ML programs. The best-fitting available model for ML analyses was LG + I + F with class weights optimized from the data set. We used this model to estimate the “posterior mean site frequencies” using the PHYLLOBAYES tree as a guide tree (using the exchangeabilities from the LG matrix; Wang et al. 2014; Pânek et al. 2016) followed by tree-searching and bootstrapping (http://www.iqtree.org/doc/Complex-Models), last accessed December 20, 2016). Topological support for trees of the supermatrix was conducted using 1,000 MLBS pseudoreplicates (fig. 3, see supplementary fig. S1, Supplementary Material online).

Data Availability
All transcriptomic data generated in this manuscript have been deposited with National Center for Biotechnology Information under the BioProject PRJNA380424 as detailed in supplementary table S2, Supplementary Material online. All single gene alignments, masked and unmasked, and phylogenomic matrices are available in the supplementary file Kang_etal.2017.tar.gz.

Supplementary Material
Supplementary data are available at Molecular Biology and Evolution online.

Author Contributions
M.W.B., F.W.S., A.K., and D.J.G.L. conceived this project and experiments. S.K., M.W.B., A.K.T., J.D.S., I.C., T.P., D.M.C.A, A.K., D.J.G.L., M.K., L.L.S., and F.W.S. collected, isolated, and maintained cultures of the amoebozoans sampled within. S.K., A.K.T., D.J.G.L., I.C., T.P., and M.W.B. collected transcriptomic data. M.W.B. and A.J.R. conceived phylogenomic experiments. S.K., A.K.T., and M.W.B. conducted the phylogenomic experiments. A.S. and A.K. helped with in depth interpretation of results. S.K. composed the first draft of this manuscript. All authors contributed to and agree with the text within the manuscript.

Acknowledgments
Mississippi State University’s High Performance Computing Collaboratory provided some computational resources. We wish to thank the anonymous reviewers of this manuscript who provided detailed constructive comments. We wish to thank Dr Franck Gael Carbonero at the University of Arkansas for the MiSeq sequencing. Additionally, we thank Génome Québec Innovation Centre for HiSeq sequencing. This project was supported by the National Science Foundation (NSF) Division of Environmental Biology (DEB) grant 1456054 (http://www.nsf.gov), awarded to M.W.B., F.W.S., and A.K., the São Paulo Research Foundation (FAPESP) Young Investigator Award to D.J.G.L. (#2013/04585), and the Henry Hunter Family Research Foundation at Mississippi State University Initiation Award to M.W.B. A.J.R. thanks the Canada Research Chairs Program for support.

References
Adl SM, Simpson AG, Lane CE, Lukes J, Bass D, Bowser SS, Brown MW, Burki F, Dunthorn M, Hampf V, et al. 2012. The revised classification of eukaryotes. J Eukaryot Microbiol. 59:429–493.
Amaral Zettler LA, Neral TA, O’Kelly CJ, Peglar MT, Gillevet PM, Bolivar I, Fahrni JF, Silverman JD, Sogin ML. 2000. A molecular reassessment of the Leptomyxid amoebae. Protist 151:275–282.
Berney C, Geisen S, Van Wych J, Nitsche F, Vanormelingen P, Bonkowski M, Bass D. 2015. Expansion of the ‘reticulosphere’: diversity of novel branching and network-forming amoebae helps to define Varicosia (Amoebozoa). Protist 166:271–295.
Bolivar I, Fahrni JF, Smirnov AV, Pawlowski J. 2001. LSU rRNA-based phylogenetic position of the genera Amoeboidea and Chaos (Lobosea, Gymnamoebia); the origin of gymnamoebae revisited. Mol Biol Evol. 18:2306–2314.
Brown MW, Sharpe SC, Silverman JD, Heiss AA, Lang BF, Simpson AGB, Roger AJ. 2013. Phylogenomics demonstrates that breviate flagellates are related to opisthokonts and apusomonads. Proc R Soc B: Biol Sci. 280:20131755–20131755.
Brown MW, Silverman JD. 2013. The non-dictyostelid sorocarpic amoebae. In Romeralo, Escalante, Baldauf, editors. Dictyostelids – evolution, genomics and cell biology. Heidelberg (Germany): Springer. pp 219–242.
Brown MW, Silverman JD, Spiegel FW. 2011. “Slime molds” among the Tubulinea (Amoebozoa); molecular systematics and taxonomy of Capromyxa. Protist 162:277–287.
Brown MW, Spiegel FW, Silverman JD. 2007. Amoeba at attention: phylogenetic affinity of Sappinia pedata. J Eukaryot Microbiol. 54:511–519.
Cavalier-Smith T. 1998. A revised six-kingdom system of life. Biol Rev Cambridge Philos Soc. 73:203–266.
Cavalier-Smith T, Chao EE, Lewis R. 2016. 187-Gene phylogeny of protozoan phylum Amoebozoa reveals a new class (Cutosoa) of deep-branching, ultrastructurally unique, enveloped marine Lobosa and clarifies amoeba evolution. Mol Phylogenet Evol. 99:275–296.
Cavalier-Smith T, Chao EEY, Oates B. 2004. Molecular phylogeny of Amoebozoa and the evolutionary significance of the unikont Phalasterium. Eur J Protistol. 40:21–48.
Cavalier-Smith T, Fiore-Donno AM, Chao E, Kudryavtsev A, Berney C, Snell EA, Lewis R. 2015. Multigene phylogeny resolves deep branching of Amoebozoa. Mol Phylogenet Evol. 83:293–304.
Eichinger L, Pachebat JA, Glickner G, Rajandream MA, Sucgang R, Berriman M, Song J, Olsen R, Sazhfranski K, Xu Q, et al. 2005. The genome of the social amoeba Dictyostelium discoideum. Nature 435:43–57.
Eme L, Sharpe SC, Brown MW, Roger AJ. 2014. On the age of eukaryotes: evaluating evidence from fossils and molecular clocks. Cold Spring Harb Perspect Biol. 6.
Fahrni JF, Bolivar I, Berney C, Nasserova E, Smirnov AV, Pawlowski J. 2003. Phylogeny of lobose amoebae based on actin and small-subunit ribosomal RNA genes. Mol Biol Evol. 20:1881–1886.
Fiore-Donno AM, Nikolaev SI, Nelson M, Pawlowski J, Cavalier-Smith T, Baldauf SL. 2010. Deep phylogeny and evolution of slime moulds (mycetozoa). Protist 161:55–70.
Goodenough U, Heitman J. 2014. Origins of eukaryotic sexual reproduction. Cold Spring Harb Perspect Biol. 6:a016154–a016154.
Jeffroy O, Brinkmann H, Deluc F, Philippe H. 2006. Phylogenomics: the beginning of incongruence? Trends Genet. 22:225–231.
Kudryavtsev A, Brown MW, Tice A, Spiegel FW, Pawlowski J, Anderson OR. 2014. A revision of the order pellitida Smirnov et al., 2011 (Amoebozoa, Discosea) based on ultrastructural and molecular evidence, with description of Endostelium crystalliferum n. sp. Protist 165:208–229.
mandrillaris, Naegleria fowleri, and Sappinia diploidea. FEMS Immunol Med Microbiol. 50:1–26.
Walker G, Simpson AGB, Edgcomb V, Sogin ML, Patterson DJ. 2001. Ultrastructural identities of Mastigamoeba punctachora, Mastigamoeba simplex and Mastigella commutans and assessment of hypotheses of relatedness of the pelobionts (Protista). Eur J Protistol. 37:25–49.
Wang HC, Susko E, Roger AJ. 2014. An amino acid substitution-selection model adjusts residue fitness to improve phylogenetic estimation. Mol Biol Evol. 31:779–792.

Whitney KD, Bennett WE, Olive LS. 1982. Observations on Echinostelium bisporum. Mycologia 74:677–680.
Yubuki N, Leander BS. 2013. Evolution of microtubule organizing centers across the tree of eukaryotes. Plant J. 75:230–244.
Zadrobilkova E, Walker G, Cepicka I. 2015. Morphological and molecular evidence support a close relationship between the free-living archamoebae Mastigella and Pelomyxa. Protist 166:14–41.