doi 10.18699/bgrs2024-5.2-26

Single-cell genomics – a powerful tool to resolve the phylogeny of uncultivable species of Amoebozoa

Smirnov A.[1]*, Mesentsev E.[1], Kamyshatskaya O.[1, 2], Glotova A.[1, 2], Bondarenko N.[1], Rayko M.[2], Skalon E.[1], Lotonin K.[1], Nassonova E.[1, 2]

[1] *Department of Invertebrate Zoology, Faculty of Biology, Saint Petersburg University, St. Petersburg, Russia*

[2] *Laboratory of Cytology of Unicellular Organisms, Institute of Cytology, RAS, St. Petersburg, Russia*

** alexey.smirnov@spbu.ru*

Key words: Amoebozoa; Tubulinea; Discosea; Variosea; genomics; phylogeny; sequencing

Motivation and Aim: Amoebozoa still remain one of the least studied lineages of protists. Unstable body shape, difficulty of cultivation and study and inability of most species to grow in culture make investigation of the diversity and phylogeny of this group very difficult, and in some cases – almost impossible. Until recently, uncultivable species of Amoebozoa were absent in phylogenetic trees an their position in the system of Amoebozoa was defined based on evidences from morphology [1]. This problem seriously embarrassed further development of amoebozoan phylogeny [2]. Many species cannot multiply in laboratory culture and are recovered only in mixed, initial cultures and in a little number of specimens. To solve the problem, several years ago we elaborated and successfully applied protocols and approaches allowing us to extract DNA from a single amoebozoan cell and to perform the whole-genome amplification of this DNA [3, 4]. Using these techniques, we successfully found home for several poorly studied lineages of Amoebozoa using both single-gene phylogeny and phylogenomic approaches. In particular, in this way we have studied a very rare species *Thecochaos fibrillosum*, not seen by investigators for almost hundred years and established its position in the tree of Amoebozoa. We complemented the amoebozoan tree with numerous species belonging to the genus *Polychaos*, previously unavailable for molecular studies and made several more interesting findings in the filed of Amoebozoa phylogeny and phylogenomics.

Methods and Algorithms: For DNA isolation, individual amoeba cells were collected from the mixed cultures manually, with tapered-tip glass Pasteur pipette, or using Eppendorf micromanipulators/micronjector mounted on Leica DMI3000 inverted microscope equipped with IMC contrast. Individual cells were transferred into 40 mm Petri dishes filled with fresh Millipore-filtered (0.22 μm) PJ medium, and left to starve for three days. Every day the medium was replaced with the fresh one. Dishes with cells were examined using daily to check for the absence of visible food vacuoles inside cells, absence of fungal contamination, and other eukaryotes in the surrounding medium. Further, individual cells were collected and placed with 1–2 μl of the medium in 200-μl PCR tubes. DNA was extracted using the Arcturus PicoPure DNA Extraction Kit (Thermo Fischer Scientific, USA). We performed the whole genome amplification of the total DNA using Multiple Displacement Amplification (MDA) with slightly modified protocol for REPLI-g Single Cell DNA Amplification Kit (Qiagen, Hilden, Germany). The resulting MDA products were sequenced using Illumina HiSeq 2500 system. Quality control check of raw sequence data was performed using FastQC. SPAdes

assembler was used for *de novo* genome assembly [5]. ML analysis was performed by IQ-TREE v1.6.12 [6] with the model GTR+I+G+F. For multigene phylogeny, putative proteins were predicated using Metaeuk or TransDecoder (https://github.com/TransDecoder/TransDecoder/wiki). The base for the alignment was the dataset obtained by Kang et al. [7] (323 genes). Potential orthologs were identified using BLAST on protein datasets from potential closest relatives of studied organisms. ML trees were constructed for each protein-coding gene separately using FastTree [8] and examined manually to estimate the overall tree configuration and the potential presence of paralogs of the target gene. The set of positions for the analysis was selected using g-blocks as implemented in SeaView 4.0 [9]. RaxML 8.2.12 [10], and Phylobayes 1.8 [11] were run at Cipres portal [12].

Results: The above-described approach allowed us to perform several case studies and characterize a number of amoebozoan species, seen only in mixed culture and in very limited number of specimens. Until recently correct identification and record of such organisms was believed to be impossible. One of them is a member of the genus *Thecochaos* – the species *T. fibrillosum.* It was isolated from the soil of West Siberia, and we had ca. 15 cells of this organism. However, using single-cell techniques, we were able to obtain light- and electron-microscopic data, get SSU sequence and perform whole-genome amplification for NGS sequencing and build multigene phylogeny. The results of this study allowed this organism to find its home within the family Thecamoebidae (Amoebozoa, Discosea). The second case study is dedicated to very unusual amoeba with characteristic rolling movement of the cell surface. This organism was frequently seen from many habitats, but all attempts to culture it or to accumulate sufficient number of cells for DNA extraction failed. The organism is small (usually, 10–20 microns), thus single cells were collected using micromanipulation technique. This way to collect cells resulted in very clean DNA preparations. To establish its position in the phylogenetic tree of Amoebozoa we build a comprehensive multigene alignment consisting of ca. 70 000 AA positions. The resulting tree placed the studied organism in the Cutosea clade, which is congruent with LM and TEM data. The resulting Amoebozoan tree supported our hypothesis on the basal position of Tubulinea lineage in the tree of Amoebozoa [13]. In the latest studies we started to use not the whole cell, but the isolated cell nucleus to get DNA sample for the whole genome amplification. It opened a way to get sequences of largest species of Amoebida. These organisms are always filled with food vacuoles and endocytobionts. Until recently almost all attempts to get nuclear DNA from such species failed. Isolation of the nucleus allows getting pure genomic DNA and include these organisms in the phylogenetic and phylogenomic analysis [14].

Conclusion: Single-cell genomics is a powerful tool allowing phylogenetic placement of uncultivable amoeboid organisms found in the environment. This approach is especially fruitful in amoebozoan ecology, since identification of as many morphospecies as possible may be decisive for the quality of the study. Modern algorithms of data sorting and assemble allow to properly clear the target genome from contaminating reads. Using single-cell techniques, we were able to obtain light- and electron-microscopic data, get SSU sequence and perform whole-genome amplification for NGS sequencing and multigene phylogeny for several amoebae species. The number of cells necessary for a comprehensive study was reduced with these approaches from thousands to ten or less. Hence, establishing of stable culture is no more a necessary requirement for identification and proper description of amoeba species. Obtaining whole-genome data

allows direct building of phylogenomic trees, which are much more reliable rather than the single gene tree. These studies solved several old and frequently mentioned puzzles in Amoebozoan phylogeny and significantly improved the overall tree of Amoebozoa.

Funding: The study was supported by the Russian Science Foundation project 24-44-00096 and used equipment of the core facility centers “Culturing of microorganisms”, “Development of molecular and cell technologies” and “Biobank” of the Research Park of Saint Petersburg University.

References

1. Smirnov A., Chao E., Nassonova E., Cavalier-Smith T. A revised classification of naked lobose amoebae (Amoebozoa: Lobosa). *Protist.* 2011;162:545-570. doi 10.1016/j.protis.2011.04.004
2. Adl S.M., Bass D., Lane C.E., Lukeš J., Schoch C.L., Smirnov A. et al. Revisions to the classification, nomenclature, and diversity of Eukaryotes. *J Eukaryot Microbiol*. 2018;66:4-119. doi 10.1111/jeu.12691
3. Mesentsev Y., Bondarenko N., Kamyshatskaya O. et al. Thecochaos is not a myth: study of the genus Thecochaos (Amoebozoa, Discosea) – a rediscovered group of lobose amoeba, with short SSU gene. *Org Divers Evol.* 2023;23:7-24. doi 10.1007/s13127-022-00581-9
4. Kamyshatskaya O., Nassonova E., Smirnov A. Phylogenetic position of amoeba genus *Deuteramoeba* (Amoebozoa, Tubulinea). *Protistology.* 2017;11:231-237. doi 10.21685/1680-0826-2017-11-4-4
5. Nurk S., Bankevich A., Antipov D. et al. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. *J Comput Biol*. 2013;20:714-737. doi 10.1089/cmb.2013.0084
6. Nguyen L., Schmidt H., von Haeseler A., Minh B. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 2015;32(1):268-274. doi 10.1093/molbev/msu300
7. Kang S., Tice A.K., Spiegel F.W., Silberman J.D. et al. Between a pod and a hard test: the deep evolution of amoebae. *Mol Biol Evol.* 2017;34:2258-2270. doi 10.1093/molbev/msx162
8. Price M.N., Dehal P.S., Arkin A.P. Fasttree: Computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol.* 2009;26:1641-1650. doi 10.1093/molbev/msp077
9. Gouy M., Guindon S., Gascuel O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol*. 2010;27(2):221-224. doi 10.1093/molbev/msp259
10. Stamatakis A. RAxML Version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 2014;30:1312-1323
11. Lartillot N., Lepage T., Blanquart S.I. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics.* 2009;25:2286-2288
12. Miller M.A., Pfeiffer W., Schwartz T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Gateway Computing Environments Workshop (GCE) IEEE. 2010;1-8
13. Tekle Y.I., Wang F., Wood F.C., Anderson O.R., Smirnov A. New insights on the evolutionary relationships between the major lineages of Amoebozoa. *Sci Rep.* 2022;12:11173. doi 10.1038/s41598-022-15372-7
14. Kamyshatskaya O.G., Bondarenko N.I., Mesentsev Y.S., Chistyakova L.V., Nassonova E.S., Smirnov A.V. Molecular phylogeny of Polychaos annulatum (Amoebozoa, Tubulinea, Euamoebida) shows that genus Polychaos belongs to the family Hartmannellidae. *J Eukaryot Microbiol*. 2020;67:321-326. doi 10.1111/jeu.12782