GIULIA MAGRI RIBEIRO

Sequenciamento e anotação do transcriptoma da ameba tecada *Arcella intermedia*: Descrição de vias e descobertas de genes.

Transcriptome sequencing and annotation of the testate amoeba *Arcella intermedia*: Pathway description and gene discovery.

São Paulo
2018

GIULIA MAGRI RIBEIRO

Sequenciamento e anotação do transcriptoma da ameba tecada *Arcella intermedia*: Descrição de vias e descobertas de genes.

Transcriptome sequencing and annotation of the testate amoeba *Arcella intermedia*: Pathway description and gene discovery.

Dissertação apresentada ao Instituto de Biociências da Universidade de São Paulo para obtenção do título de Mestre em Zoologia pelo Programa de Pós-graduação em Zoologia.

Versão corrigida contendo as alterações solicitadas pela comissão julgadora em 30 de Outubro de 2018. A versão original encontra-se no Instituto de Biociências da USP e na Biblioteca Digital de Teses e Dissertações da USP.

Supervisor: Prof. Dr. Daniel J.G. Lahr

São Paulo

2018

Ficha catalográfica elaborada pelo Serviço de Biblioteca do Instituto de Biociências da USP,
com os dados fornecidos pelo autor no formulário:
http://www.ib.usp.br/biblioteca/ficha-catalografica/ficha.php

Ribeiro, Giulia Magri
Sequenciamento e anotação do transcriptoma da ameba tecada Arcella intermedia: Descrição de vias e descobertas de genes. / Giulia Magri Ribeiro; orientador Daniel José Galafasse Lahr. -- São Paulo, 2018.

118 f.

Dissertação (Mestrado) - Instituto de Biociências da Universidade de São Paulo, Departamento de Zoologia.

1. Teca ameba. 2. Transcriptoma. 3. Transferências laterais de genes. 4. Metabolismo anaeróbio. I. Lahr, Daniel José Galafasse, orient. II. Título.

Bibliotecária responsável pela estrutura da catalogação da publicação:
Elisabete da Cruz Neves – CRB – 8/6228

Comissão julgadora:

_____

Prof. Dr. _____

_____

Prof. Dr. _____

_____

Prof. Dr. _____

Orientador

Resumo

RIBEIRO, Giulia Magri. **Título do trabalho**: Sequenciamento e anotação do transcriptoma da ameba tecada *Arcella intermedia*: Descrição de vias e descobertas de genes. 2018. 117 f. Dissertação (Mestrado em Zoologia) – Instituto de Biociências, Universidade de São Paulo, São Paulo, 2018.

Arcella Ehrenberg 1832 é um dos gêneros de tecamebas mais numerosos, pertencente aos Arcellinida. Estas são uma linhagem aeróbia de amebas tecadas que vivem em uma grande variedade de ambientes. Provavelmente, sua capacidade de sobreviver em condições tão divergentes está relacionada a algum grau de flexibilidade metabólica. Os organismos anaeróbicos ganharam e perderam vários genes relacionados ao metabolismo energético. Este processo modifica mitocôndrias clássicas até a perda da função e transformação em organelas relacionadas (mitossomos e hidrogenossomos). Aqui proponho que a adaptação de *Arcella intermedia* a ambientes microaerófilos está relacionada à aquisição de novos genes. Existem dois modos principais de aquisição de genes. Na visão tradicional, a duplicação gênica é responsável por gerar diversidade, seguida por mutações e neofuncionalidade da duplicata. Alternativamente, os genes podem ser adquiridos de outras espécies (transferências laterais de genes). O segundo processo tem uma grande importância evolutiva e é ainda pouco considerado na evolução eucariótica. Por isso, também proponho neste trabalho que genes relacionados ao metabolismo anaeróbico em *Arcella* sejam adquiridos por transferência lateral de genes. Entretanto, a análise de dados genômicos e transcriptômicos é inexistente para *A.intermedia*. A caracterização de dados em escala genômica de eucariotos é essencial para a descoberta de genes e para a inferência transições sobre a árvore da vida. O conjunto de dados de transcriptoma deste trabalho fornece um primeiro esforço de caracterização de sequências expressas em *A. intermedia*. Utilizamos extrações de célula-única em diferentes momentos de crescimento e extração de RNA de cultura inteira, a fim de aumentar a diversidade de momentos metabólicos das células. Sequências mapeadas permitiram identificar vias funcionais em células de *A. intermedia*. Em geral, parece que genes relacionados a processos metabólicos são os que aparecem mais frequentemente, seguidos dos de sinalização e respostas a estímulos. Nós descrevemos a função do metabolismo de carboidratos e energia, incluindo uma via anaeróbica. Encontramos em *A.intermedia* os genes ACS-ADP e PFO. Descrevemos o metabolismo de aminoácidos, com pelo menos 12 vias metabólicas de aminoácidos descritas e catabolismo relacionado a intermediários do ciclo de TCA. Cálcio, Ras GTPases, PI3K-AK e AMPK-mTOR são as principais vias de sinalização representadas nos transcriptomas. Descrevemos importantes vias para amebas, que são

endocitose e fagocitose. Parecem ser vias semelhantes àquelas já descritas para outras amebas, com dependência de F-actina e pequenas GTPases da subfamília Rho. Não conseguimos encontrar muitas informações sobre a morte celular programada em *A. intermedia*, mas o crescimento celular é semelhante com as vias descritas para os dinoflagelados. Esperamos que os próximos genomas terminem a descrição da função desses organismos, mas acreditamos que nosso trabalho já é um bom ponto de partida. A fim de obter uma visão mais clara da presença de genes de metabolismo anaeróbico em Amoebozoa, realizamos buscas no BLAST em bancos de dados de Amoebozoa e Arcellinida, para a presença/ausência de ACS-ADP, PFO e [FeFe] -H2ase. Outras espécies de Arcellinida também apresentaram estes genes, *Difflugia sp*, *Difflugia compressa* e *Cyclopyxis lobostoma*. Além destes, os já conhecidos *Mastigamoeba balamuthi*, *Entamoeba histolytica* e *Acanthamoeba castelanii*. Sequências de amebozoários não formam um grupo monofilético em nenhum dos três genes. No entanto, as sequencias de Arcellinida sempre se agrupam. Como são grupos de Amoebozoa de tal maneira distintos que possuem estes genes de metabolismo anaeróbico, e sendo que a maioria não possui, é mais provável que sejam transferências laterais independentes entre esses grupos de ameba, gerando a possibilidade de ocupar um novo nicho. O objetivo principal deste trabalho foi gerar ferramentas para entender a capacidade de algumas amebas tecadas em resistir a condições adversas do meio ambiente. Encontramos muitas questões interessantes, mas a que teve nosso foco nesta dissertação foi (1) a evolução de genes relacionados ao metabolismo anaeróbio em linhagens de amebas tecadas. Os dados da sequência reunidos e anotados estarão disponíveis como sequências de referência, facilitando o trabalho com esse grupo. Os resultados também podem ser aplicados aos marcadores de biomonitoramento para o gerenciamento dos recursos hídricos. Este trabalho irá melhorar o conhecimento geral sobre a evolução e função de organismos de água doce. Esperamos também contribuir para a compreensão do impacto das transferências laterais na diversidade de Arcellinida. Palavras-chave: Tecamebas. Transcriptoma. Transferências laterais de genes. Metabolismo anaeróbico.

Abstract

RIBEIRO, GIULIA MAGRI. **Work title**: Transcriptome sequencing and annotation of the testate amoeba *Arcella intermedia*: Pathway description and gene discovery. 2018. 117 p. Dissertation (Master of Zoology) – Biosciences Institute, University of Sao Paulo, Sao Paulo, 2018.

Arcella Ehrenberg 1832, is a diverse genus of testate amoeba with more than 130 taxa described and still underestimated. Arcellinids are an aerobic lineage of testate amoeba that lives in a wide variety of environments. Probably their ability to survive in such divergent conditions is related to some degree of metabolic flexibility. Anaerobic organisms have gained and lost a number of genes related to energetic metabolism. These processes modify classic mitochondria until loss of function and transformation in mitochondrial related organelles (mitosomes and hydrogenosomes). Here I propose that *Arcella intermedia* adaptation to microarophilic environments are related to the acquisition of new genes. There are two main modes of acquisition of new genes – the traditional view, where duplication is followed by mutations and neo-functionality of the duplicate. Or genes can be acquired from other species (lateral gene transfers). The second process has a major importance in prokaryotic evolution and is probably under considered in eukaryotic evolution. I also propose in this work that genes related to anaerobic metabolism in *Arcella* are acquired by lateral gene transfer. However, analysis of genomic and transcriptomic data are absent for *A.intermedia*. The transcriptome dataset from this work provides the first effort of characterization of expressed sequences in *A.intermedia*. We used a single cell from different moments of growth and whole culture RNA extraction in order to increase the diversity of metabolic moments of the cells. Mapped sequences allowed us to identify functional pathways in *A.intermedia* cells. In general, it seems that metabolic processes are showing up more, followed by signaling and responses to stimuli. We describe functioning of carbohydrate and energy metabolism including even an anaerobic pathway to produce energy. We found ACS-ADP and PFO genes in *A.intermedia*. We describe amino acid metabolism, with at least 12 amino acids metabolizing pathways described and catabolism mainly related to TCA cycle intermediates. Calcium, Ras GTPases, PI3K-AK and AMPK-mTOR are the main signlaing pathways represented in transcriptomes. We described important pathways for amoeba: endocytosis and phagocytosis and it seems to be similar with the ones already described for other amoebae with a high abundance on F-actin and small GTPases of Rho subfamily. We couldn't find lots of information about programmed cell death in *A.intermedia*, however cell growth is similar to pathways described for dinoflagellates. We expect

that upcoming genomes will finish the description of the functioning of those organisms, but we believe our work already is a good starting point. To visualize and understand the presence of anaerobic metabolism genes in Amoebozoa, we conducted BLAST searches in Amoebozoa and Arcellinida data bases for the presence/absence of ACS-ADP, PFO and [FeFe]-H2ase. Other Arcellinida species also presented these genes, *Difflugia sp.*, *Difflugia compressa* and *Cyclopyxis lobostoma*. Besides these, the already known *Mastigamoeba balamuthi*, *Entamoeba histolytica* and *Acanthamoeba castelanii*. Amoebozoa sequences don't form a monophyletic group in any of the three genes. However, Arcellinida sequences always grouped. As such distinct amoeba groups have those anaerobic metabolism genes, however, most of the Amoebozoa do not. It is more likely to think of lateral transfers occurring independently among these amoeba groups, generating the possibility of occupying a new niche. The main objective of this work was to start generating tools to understand the ability of some testate amoeba to resist harsh environmental conditions. We found lots of interesting questions but the one we focused on this dissertation was (1) the evolution of anaerobic related genes in testate amoeba lineages. The assembled and annotated sequence data will be available as reference sequences, making the work with this group easier. The results can also potentially be applied as biomonitoring markers for the management of water resources. This work will improve the general knowledge on the evolution and function of freshwater organisms. We also expect to contribute to the understanding of the impact of lateral gene transfers in Arcellinida diversity.

Key words: Testate-amoeba. Transcriptome. Lateral gene transfers. Anaerobic metabolism.

0.1 Organization of the dissertation

This dissertation has 4 chapters. The main chapters are the two manuscripts produced (Chapters 2-3) and your respective supplementary files (Appendix B and C). It also has a general introduction (Chapter 1) with the general methodology detailed and a general discussion and conclusion (Chapter 4) summarizing the main results of the dissertation. Chapter 2 and 3 are being prepared to publish. The intended articles, Chapter 2 and 3, will be submitted to the Journal of Eukaryotic Microbiology and Journal Molecular Biology and Evolution, respectively.

0.2 General considerations

In this work, I characterize genomic-scale data from an important species of testate amoeba, *Arcella intermedia*, and I demonstrate genes related to metabolic novelties. I used transcriptomic high throughput sequencing techniques as the first approach to this characterization. I combined the experimental acquisition of data with bioinformatics work using data available on public databases. Genes of interest were processed and analyzed using phylogenetic techniques. I approached these goals from the perspective that evolutionary novelties may be related to the acquisition of new genes and pathways. The idea of "niche" is important in this work because species often experience phenotypic barriers to colonize a new environment. *Arcella intermedia* is a group of testate amoeba that seems to have overcome several of those barriers.

Chapter 1, is a general introduction where I am going to try to bring some essential aspects related to the problematics of work. I am going to start making an approximation of the state of the art in eukaryotic microorganisms studies nowadays. I will explain the environmental conditions of laboratory growth, how

metabolism works and how microbial eukaryotes can adapt do oxygen-limited conditions. I also am going to start making an introduction to *Arcella intermedia* the focus group of my study. *Arcella*, which occupies a wide variety of environments. Probably acquired common adaptations to those environments during evolutionary time. How genes related to adaptations are acquired is an interest of the study of genetic organization and evolution and how "big data" technologies can be used to address evolutionary and ecological questions are all introduced before the presentation of the general methodology.

Chapter 2, "*De novo* sequencing, assembly and annotation of transcriptome for the free-living testate amoeba *A.intermedia*", has importance as the first description of the transcriptome of a testate amoeba. I describe the entire process of obtaining a transcriptome, processing data and annotation. I also compare different methods of analysis. My main focus on this work was also to present the main metabolic processes developed for this organism. Additionally, this work reveals the probable importance of calcium and RAS signaling pathways in *A.intermedia*. In general metabolism, I demonstrate the probable appearance of the anaerobic pathways in this group. In cell growth, I describe the similarity of function between *A.intermedia* and dinoflagellates. This work drove the focus of chapter 3, "Phylogenomics reveals potential low-oxygen adaptations in free-living testate amoebae".

In chapter 3, I investigate the origin of 3 genes related to anaerobic metabolism in some Arcellinida. I hypothesized that these genes could be laterally transfered independently in some groups. In chapter 3 I show the presence and absence of anaerobic metabolism genes, their historical reconstruction (trees) in an attempt to understand their evolutionary history.

0.3 Contribution

Taken together, these chapters begin to tell the history of the evolution of new characteristics in Arcellinida through the acquisition of genes and metabolic pathways. We still need to understand, for example, in what scenario these pathways are important; how they are regulated; how exactly those genes were acquired, and; what is the impact of these acquisitions and metabolic organizations on the diversity of groups. In this work we start to address some of these questions and provide some possible answers: for example, genes seem to be acquired independently for organisms by lateral gene transfers and the fixation or not of the new characteristic maybe gave opportunity for the species to occupy the niche which they currently inhabit.

1 General Introduction

1.1 Studying microbial eukaryotes

Free-living microbial eukaryotes are a group of mostly unicellular eukaryotic organisms, commonly called protists. Protists are extremely abundant and present in most terrestrial environments. They compose most of the global eukaryotic diversity (MEDINGER et al., 2010; PAWLOWSKI et al., 2012; BOENIGK et al., 2012) (Figure 1.1).

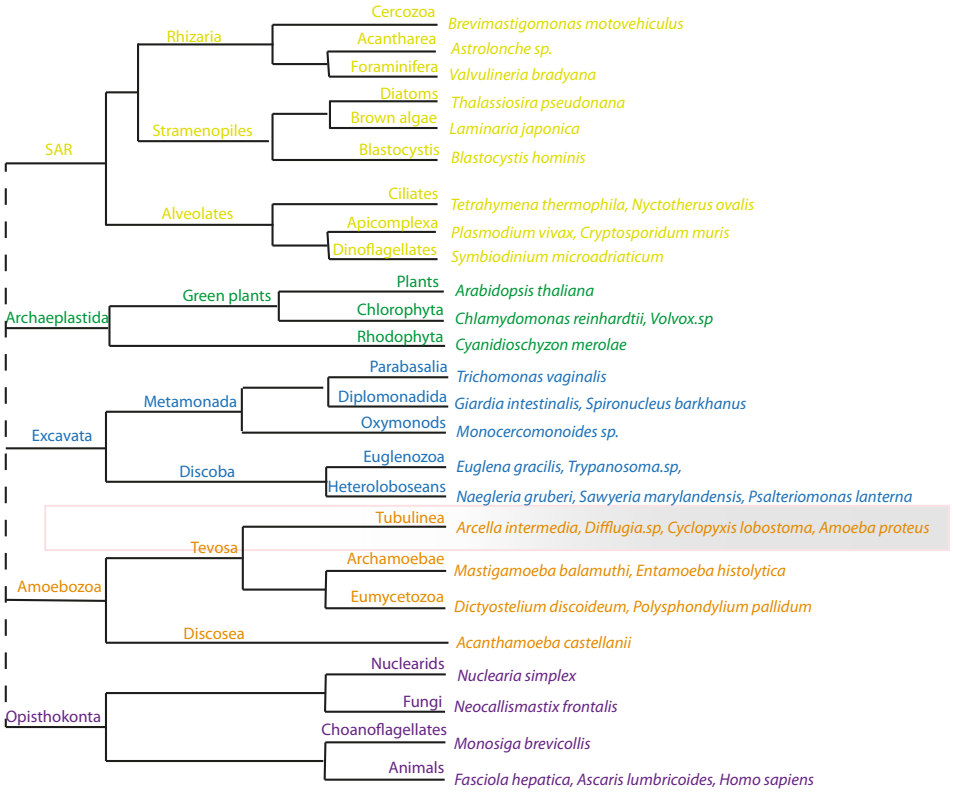
The study of protist diversity and functional roles has gained importance among researchers (CORLISS, 2002; KEELING; CAMPO, 2017). The fields with greater focus are: their use for environmental monitoring, taxonomy and phylogenetic relationships between groups, forensic biology and paleontology (ADL; GUPTA, 2006; WILKINSON; CREEVY; VALENTINE, 2012; SZELECZ et al., 2014; PAWLOWSKI et al., 2014). However, current knowledge about the most basic aspects of diversity, systematics, physiology and genetics of protists is still very limited. Some of the limitations of the area are: insufficient sampling; difficulty in observation and cultivation; lack of diagnostic characters or even the lack of a developed methodology of work with these organisms (ŠLAPETA; MOREIRA; LÓPEZ-GARCÍA, 2005).

1.2 Microbial ecology

1.2.1 Population Growth

Microbial growth is a process that much of our knowledge, comes from controlled laboratory studies and is characterized by a metabolic preparation that results in cell division (MAIER; PEPPER; GERBA, 2009). Much of our knowledge

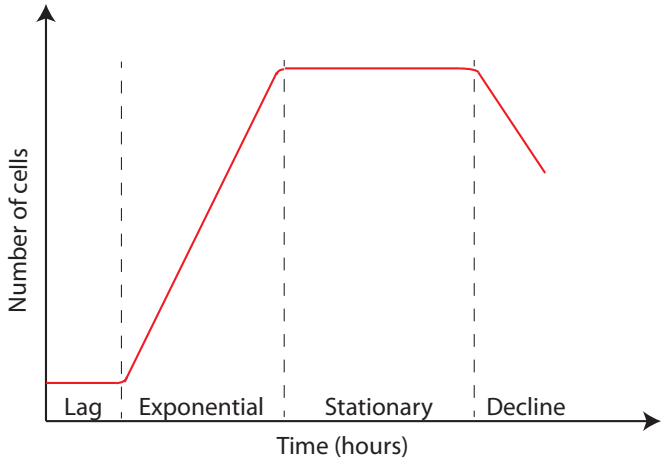
Figure 1.1 – Illustration of the eukaryotic tree of life. The tree represents the current interpretation of the relations between the supergroups of eukaryotes. The studied lineage is highlighted in red and shaded.



about the growth of microorganisms results from controlled laboratory studies using pure bacterial cultures. Typical growth curves have 4 phases: lag (no growth, adaptation to environment and mass gain), log (exponential growth), stationary (equilibrium between division and death) and decline (reduction of population density) (Figure 1.2).

A cell culture must comprise all the conditions necessary for the cell to survive and to proliferate (CASTILHO et al., 2008). In cellular cultures, the intensity

Figure 1.2 – Example of a population growth curve model. The different phases of the curve are highlighted.



of the cell growth is determined by several factors: pH, osmolarity, temperature, the concentration of essential gases (oxygen and CO_2), available substrate, and even the state of the cell at the time of inoculation (FRESHNEY, 1994). Inadequate culture conditions result in decreased cell viability. Some factors are determinant in limiting the growth of cells in culture, such as nutrient shortage, accumulated metabolites, decreased dissolved oxygen level and lack of surface for cell adhesion.

- **Metabolic accumulation**: Most cell lineages can tolerate some variation in osmolarity in relation to the medium. The development of cells in culture can also modify this osmolarity. For example, an increase in CO_2, with cellular respiration, tends to increase the acidification of the medium (CASTILHO et al., 2008). From physical processes, osmolarity may also be increased by evaporation, i.e. the culture flasks are not fully sealed and allow a balance between the culture medium and air CO_2.
- **Oxygen content**: Oxygen is the most important component for cellular respiration and probably to the cell viability. Therefore, with the culture

time the oxygen tends to become limiting (CASTILHO et al., 2008). This is because oxygen is poorly soluble in aqueous media.

- **Nutritional limitation**: Cells need a range of essential nutrients to survive (CASTILHO et al., 2008). For the cell to survive and proliferate, it needs to attend your metabolic requirements properly. When the cell fails to attend its own energy demands, it tends to enter in apoptosis or another cellular death process (AL-RUBEAI; SINGH, 1998).

1.2.2 Microbial metabolism

Every living cell has "chores". These "chores" consist of synthesis activity or functional activity. Synthesis activity, for example, is the production of cytoplasm and its components. Functional activity, for example, is contraction or locomotion. For any of these "chores" the cell needs energy. In contrast to the plant cell, that is can produce your own energy with the photosynthesis, animal cells get its energy through the breakdown of nutrients obtained from food sources, mainly carbohydrates and aminoacids.

The degradation of glucose and other substrates is made by a series of small reactions. Enzymes mediate each of these reactions. They all require oxygen to happen, which makes it an essential component for the cellular environment. The cell can modulate the functioning of your metabolism, especially in relation to nutrition (CASTILHO et al., 2008). The co-enzyme indicative of the energy level of the cell is the adenosine triphosphate (ATP). The catabolism of substrates, such as carbohydrates, generates the production of ATP and reduced co-enzymes (NADH), both essential for cell viability. These co-enzymes are used throughout the energy metabolism for the production of other substances or reserves (HAUSER; WAGNER, 1997).

Protists, fungi and animals may vary in relation to the expected energetic metabolism, for example for an aerobic mammalian cell growing under ideal conditions. It may vary in relation to the location of enzymes, as in kinetoplastids (*Trypanosoma brucei*), that some glycolytic enzymes are found in peroxisomes (GINGER et al., 2010). Other differences may occur in the basic composition of the metabolic pathways, as in organisms that can regulate the mode of ATP production depending on the environmental stimuli received. For example, in anaerobic metabolism from *Trichomonas vaginalis*, *Entamoeba histolytica* and *Giardia lambia*. These organisms use enzymes typical of anaerobic organisms, such as pyruvate:ferrodoxin oxidoreductase(PFO) and hydrogenases (ELLIS et al., 1993; RODRÍGUEZ et al., 1996; WILLIAMS; LOWE; LEADLAY, 1987).

1.2.3 Microbial eukaryotes in oxygen-limited conditions

Life in low oxygen environments is relatively common in our planet (CATALANOTTI et al., 2013). Hypoxia may occur temporarily, as a result of geochemical, physical or biological conditions: during floods, winter ice encasements, or with high bacterial activity (eutrophication) (SIKDAR, 2017). Several protists from different groups of eukaryotes can live in oxygen-poor niches or at least survive for a limited period of time (CATALANOTTI et al., 2013; REINHARD et al., 2016).

Metabolic adaptations to anaerobiosis are usually associated with loss of mitochondrial functions (NỲVLTOVÁ et al., 2015). These functions are generally related to aerobic metabolisms such as Krebs cycle and oxidative phosphorylation. Anaerobic forms of mitochondria generate ATP exclusively by phosphorylation at the substrate level. There are several subtypes of organelles that consist of modified mitochondria: hydrogenosomes, mitosomes and anaerobic mitochondria. Aerobic respiration is the main function of mitochondria and their energy metabolism

oxidizes the molecules through the oxygen dependent electron transport chain (TSAOUSIS et al., 2012), generation ATP via ATP synthase.

Eukaryotic organisms diverged from prokaryotes in an evolutionary process of nearly 2 billion years (KURLAND; ANDERSSON, 2000). One of the major differences between eukaryotes and prokaryotes is the compartmentalization of various metabolic functions forming organelles (FUERST, 2010). One of these compartments is the mitochondria. The prevalent hypothesis of mitochondrial origin is the theory of endosymbiosis (SCHIMPER, 1883; SAGAN, 1967). The prevailing hypothesis about the origin of the mitochondria is the theory of endosymbiosis that suggests that the appearance of the organelle through endosymbiosis between an ancestral eukaryotic host and an α-proteobacteria around 1.5 billion years ago. In this process the interdependence between the host and the symbiont was so deep that most of the symbiont genes were transferred to the host or lost. Organelle functions can continue to be encoded by their own genome, or encoded by the host and marked for transfer to the mitochondria with a signal peptide.

Anaerobic organisms have gained many genes related to anaerobic metabolism (NÝVLTOVÁ et al., 2015). The metabolism of pyruvate in aerobic organisms are made through the pyruvate dehydrogenase complex (PDC). In the absence of this complex, pyruvate-ferrodoxin oxidoreductase (PFO) or pyruvate: $NADP^+$-oxidoreductase (PNO) acts on this function under anaerobic conditions (MÜLLER et al., 2012). Some anaerobic organisms can use these electrons with hydrogenasis to form molecular hydrogen directing from $NADP^+$ (MÜLLER et al., 2012). The evolution of these mitochondrial related organelles (*MROs*) is a process with intermediate steps. The functional characterization of intermediate organelles illuminates the evolutionary transitions between organelle types.

The Amoebozoa group includes both members that have or not ATP-producing organelles. *Entamoeba histolytica*, a human intestinal pathogen, have

only organelles known as mitosomes which do not produce any energy. In contrast, *Dictyostelium discoideum* is a free-living amoeba that inhabits soil and compost and possesses an aerobic mitochondria. Arcellinida are an aerobic testate amoeba that lives in a wide variety of environments. In lakes, there is a correlation between the microbiota and some parameters including the availability of oxygen. What makes these organisms capable of surviving in such divergent conditions must be related to some degree of metabolic flexibility.

1.3 Microbial genetic organization and evolution

Eukaryotic genomes have unique characteristics distinct from the prokaryotes (BROWN, 2002). Prokaryotic genomes are small in relation to eukaryotes. The physical organization is also different between eukaryotes and prokaryotes. In general, prokaryotic genomes are circular associated with smaller, independent molecules, called plasmids. In contrast, eukaryotic genomes are organized into large nuclear genomes and smaller, generally circular, mitochondrial genomes. In plants, there is a third genome, located in chloroplasts. Chromosomes, linear DNA molecules, are the primary component of the nuclear genomes. Unlike prokaryotic genomes, eukaryotes have introns in their gene organization. Eukaryotic genes also have repeated gene families with a larger number of copies than prokaryotes.

The emergence of genes generating new functions is considered a major contributor to adaptive evolutionary innovation (KAESSMANN, 2010). There are two main modes of acquiring new genes – The traditional view, with duplication followed by mutations and neo-functionality of the duplicate (KAESSMANN, 2010); Or genes can be acquired from other species (lateral gene transfers) (BROWN, 2002). However, important contributions of lateral gene transfers in the evolution of the genomes are probably still being poorly considered, especially concerning

the evolution of prokaryotes (BROWN, 2002). Implications of lateral gene transfers in prokaryotes are so high that it is impossible to study their evolutionary histories without considering horizontal flows between species (KOONIN; MAKAROVA; ARAVIND, 2001). For eukaryotic genomes, recent works are showing that apparently, the lateral gene transfer from prokaryotes occurs and that genes thus acquired through this process must have contributed significantly in the evolution of the eukaryotes (SCHÖNKNECHT; WEBER; LERCHER, 2014).

1.3.1 Study of the genetic organization in microbial eukaryotes

Phylogenetic analyses using improved methods confirmed the previous hypothesis of multiple origins of the genes in eukaryotes, being able to infer even higher numbers of gene transfers (OLSON et al., 2006). Environmental rRNA sequencing has revealed a great diversity of new protist strains. Gene expression studies gave rise to some knowledge about the response to environmental stimuli, competition, predation, and symbiotic interactions in protists. However, with their large genome sizes, they are neglected in comparison to other microorganisms (GEISEN et al., 2018). Most studies still use metazoans, parasites and some photosynthetic organisms as models. Thus, free-living protists are still the targets of less effort in this area (GRANT et al., 2012).

RNAseq

With RNAseq, it is possible to identify genes and transcripts active in cells. Thus, constructing a gene expression profile of cells that find differentially expressed genes should show us how cells respond to environmental signals. Transcripts are very cost-effective in characterizing genomic-scale data in poorly studied organisms

that do not have a reference genome (**??**). Transcripts are interesting because they are smaller and simpler, enriched by highly expressed and conserved genes (GRANT et al., 2012). The number of transcriptome studies has grown in a variety of areas such as physiology, ecology, evolution, genetics. They have led to a significant increase in gene discovery, characterization of gene expression and development of genetic markers for various sources of interest (ZHU et al., 2016; WEBER et al., 2016; SINGH et al., 2017; RISMANI-YAZDI et al., 2011; PATNAIK et al., 2016; KIM et al., 2013; GRANT et al., 2012). Currently, data on a genomic scale have developed studies in phylogenomics and improved the understanding of the tree of life (KANG et al., 2017).

Although all transcriptome usefulness, there are some sorts of problems we have to concern about when working with transcriptomic data:

- **Incompleteness of information**: The transcriptomic source of data has a noisy and incomplete nature. Transcriptomes represent a snapshot of expression from an organism. As no organism will be expressing all their genes, at the same magnitude, at the same time, we probably are looking only to a part of their history. Despite this incompleteness, many kinds of discovery came from transcriptomic data, for example: genes expressed in different developmental stages and/or different tissues of a single or related sets of species; identification of genes implicated in physiological changes such as increased tolerance to heavy metals, salinity or disease-causing organisms; gene expression leading to changes in morphology; and changes that impact genome evolution (QIU et al., 2013; YOSHIDA et al., 2016; HILLMANN et al., 2018; SINGH et al., 2017). With transcriptome and genome-scale bioinformatics continuous development, we expect that the utility of transcriptomics will only increase in evolutionary biology and that the RNA-Seq approach will offer tremendous insights into the understanding of the ontogeny and evolution of

characters. RNA-sequencing alone provides a framework for unique analyses investigating novel transcript isoforms (isoform discovery), however, to answer if the reads belongs to the organism of interest or a contaminant, it is necessary to map to the genome.

- **Unidentified Contamination**: Cultures are typically "contaminated" with the bacteria that are food for the heterotrophic amoebae. We apply many custom python scripts that remove potential bacterial contaminants. However, we can't guarantee that we removed everything. A phylogenetic signal that these sequences probably belong to the amoebae of the work is the presence of organic relations between the sequences. In all single gene trees (ACS, PFO, and Fefe), the amoebae sequences form a group even though it is a sister of a bacterial lineage. A molecular signal of the eukaryotic nature of these sequences is the peptide signal sequence in all complete sequences.
- **Single-gene trees**: A recent study with large-scale phylogenomic data has shown that phylogenetic trees constructed from a single gene may show considerable topological variation (CASTRESANA, 2007). Genes can generate different topologies by the very nature of genetic data, which can cause signal loss and artifacts. Genes accumulate mutations, which for example in species with high mutation rate, can generate long-branch attraction. Moreover, as we work with genome and transcriptome data, our sampling is still taxonomically limited, since not all species have their genomes or transcriptomes sequenced. All these factors can change the final result of the tree, resulting in loss of the phylogenetic signal.

1.4 Diversity of testate amoeba and Arcella

The testate amoebae are a polyphyletic assemblage of eukaryotic microorganisms characterized by the presence of an outer shell structure (test). Many microbial strains can build this type of shell. In addition to members in Tubulinea (Arcellinida and Corycidia), Stramenopiles, Rhizaria, including Cercozoa and dinoflagellates also have a test (JAN, 2008; NIKOLAEV et al., 2005; CAVALIER-SMITH; CHAD, 1997; BHATTACHARYA; HELMCHEN; MELKONIAN, 1995). In Tubulinea (Amoebozoa), the test seem to have originated twice: in Arcellinida and Corycidia (KANG et al., 2017). In Arcellinida, one of the most diverse strains of amoebas tested, there are 800-2,000 species, which is probably still underestimated (RIDE et al., 1999). They are part of the Amoebozoa clade, which is a sister group of the eukaryotic supergroup of Obazoa (including animals + fungi + other microbial lineages) (BROWN et al., 2013).

Arcella Ehrenberg 1832, is a diverse genera of testate amoeba with more than 130 taxa described and still underestimated (ANDREY; YURI, 2006a). Typically, *Arcella* is a binucleated amoeba with a dome-shaped shell and an aperture situated at the end of a ventral, funnel-shaped depression. Until now, researchers of the group observed only binary fission (PCHELIN, 2010). Many species have a cosmopolitan distribution, with few examples of geographically restricted distribution (BEYENS; MEISTERFELD, 2002; FERES et al., 2016). They inhabit environments of fresh water, mosses and moist soil (DEFLANDRE, 1928; ANDREY; YURI, 2006b). *Arcella* can be considered as an omnivorous group, because it is mainly a bacterial predator, but can also ingest algae and fungi.

1.5 Arcella and the environment

Arcella species are present over a wide variety of environmental conditions (ANDERSON, 2013). Because of the few amounts of studies, confusing species nomenclature and identification difficulties (KOSAKYAN et al., 2016), it is still uncertain which species are tolerant or not to harsh conditions. However, a few works already indicated which kind of environments they can probably survive (ROE; PATTERSON; SWINDLES, 2010; SILVA et al., 2013; ROE; PATTERSON, 2014; NASSER et al., 2016; MACUMBER et al., 2014; LANSAC-TÔHA et al., 2014). For example, the indication of some kinds of heavy metal contamination is possible by the appearance of *Arcella* species (SCOTT; MEDIOLI; SCHAFER, 2007). Usually is possible to find *Arcella* species in Silver, arsenic and mercury-contaminated areas. In classic gold mining areas, most of the diversity is already ruined, but *Arcella* is still present (SILVA et al., 2013). Another interesting locality where *Arcella* appeared was in devices of mechanical sewage treated aerobically, entering the anaerobic and aerobic zones of the bioreactor (PAWLOWSKI; DUDZINSKA; PAWLOWSKI, 2010). *Arcella* is also found in environments with high phosphate content and areas with advanced nitrification in process (MIHAYLOVA; PROKOPOV; MIHALKOV,). All these harsh localities where we can find *Arcella* indicate their versatility (SILVA et al., 2013). *Arcella* does invite us to investigate their metabolic functions, to understand how they can survive in all those sorts of environments.

1.6 Motivation and objectives

The objectives of this dissertation were:

- **Objective 1**: Generating transcriptomes from a non-model testate amoeba *Arcella intermedia*.

- **Objective 2:** Assembly and annotate the transcriptome generated.
    1. Hypothesis: *A.intermedia* pathways have variations in relation to the expected from "textbook" examples with model organisms.
    2. Expected outcome: Characterize *Arcella intermedia* transcriptome and give the first step in the understanding of their metabolism.
    3. Possible outcome: *Arcella intermedia* probable adaptation to microaerophilic environments are related to the acquisition of new genes.
- **Objective 3:** Generate phylogenetic trees of anaerobic metabolism-related genes.
    1. Hypothesis: Genes involved in anaerobic metabolism may be acquired by lateral gene transfers.
    2. Expected outcome: Analysis of the distribution of genes related to anaerobic metabolism in Amoebozoa.
    3. Possible outcome: Genes of anaerobic metabolism in *A.intermedia* are acquired by lateral gene transfer.

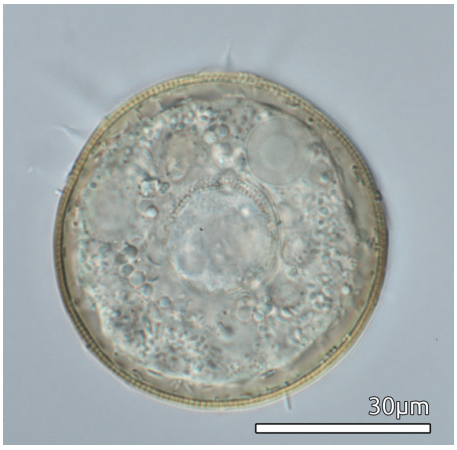
1.7 General methodology

The methodology presented in the next two chapters share several steps. Since we based all the work presented in this dissertation on transcriptomes, many steps repeat themselves (Figure 1.4). The transcriptomes of *Arcella intermedia* are a product from this work. In Chapter 2, is described more in details transcriptomic data generation, assembly, post-assembly, and annotation. In Chapter 3, is described more in details phylogenetic inference from transcriptomic data.

1.7.1 RNAseq experiment

RNAseq experiment consists of 4 main steps: Sample preparation, cDNA preparation, library preparation, and Illumina sequencing. *Arcella intermedia* used here was isolated from a small human-made pool at University of São Paulo (Coordinates - 23.565720, -46.730512). We took cells in different moments of their life in culture (LAG, LOG/Exponential, Stationary, and Decline) to increase the diversity of metabolic moments of the cells. We generated RNAseq for both single cells and whole cultures of *A.intermedia* (Figure 1.3). We followed a single cell protocol to isolate, and sequence expressed genes in *Arcella* (PICELLI et al., 2014).

Figure 1.3 – Species used in this study. A – Arcella intermedia LEP isolate 6, magnification 630X.



1.7.2 Assembly and post-assembly processing

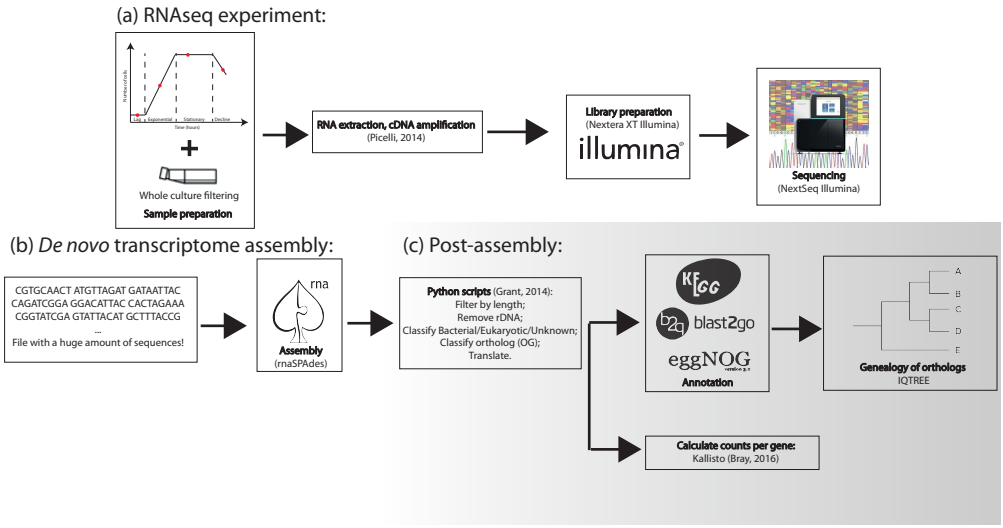
We processed raw reads for transcriptome removing low quality and adapter sequences using BBTools ⟨http://sourceforge.net/projects/bbmap⟩. After quality trimming, we used rnaSPAdes for pooling and assemble the transcriptomic

data for all samples. We processed the assembled transcriptome with a series of custom python scripts (available in ⟨http://github.com/maurerax/KatzLab/tree/HTS-Processing-PhyloGenPipeline⟩), which included a series of filtering steps for reducing contamination and processing transcriptomic data. Translated proteins were annotated using a combined strategy with EggNOG ⟨http://eggnogdb.embl.de/#/app/home⟩, Blast2GO(CONESA; GÖTZ, 2008) and BlastKOALA ⟨http://www.kegg.jp/blastkoala/⟩.

1.7.3 Genealogy generation

We used OrthoMCL ⟨http://www.orthomcl.org/orthomcl/⟩ for identification of gene families of our interest. We based the gene selection on literature and availability in the OrthoMCL database. We used HMMER (https://hmmer.org, version 3.1b2) profiles using only the conserved domains of selected genes. In Chapter 3, for example, we focused only on gene sequences from the extended glycolysis pathway. We enriched our database with Amoebozoa transcriptomic data to understand the origin and evolution of proteins of interest in this major group. Amoeba sequences were screened using the OrthoMCL seed HMMER profile of each enzyme of our interest. Non-amoebozoan lineages came from a database that consists of 603 eukaryotes, 128 archaea, 312 bacteria (GRANT; KATZ, 2014). We performed several rounds of alignments for the selected genes in Seaview (GOUY; GUINDON; GASCUEL, 2009) with alignment algorithm MAFFT (KATOH; ASIMENOS; TOH, 2009). We implemented the genealogies using maximum-likelihood optimality criteria with IQTREE.

Figure 1.4 – Scheme of reconstruction and analysis of the *Denovo* transcriptome. (a) Transcriptomic data generation consists of 4 main steps: Sample preparation: Single cells and entire culture extraction. cDNA and Illumina library preparation and Illumina sequencing. (b) Assembly of the *Denovo* transcriptome: First, the individual samples are preprocessed to remove adapters from sequences and assembled into *Denovo* contigs. (c) Post-assembly: In this step, the pre-processed samples pass through a series of python scripts for size filtering (200 bp), remove rRNA transcripts, remove bacterial identified transcripts, find the possible identity of the remaining sequences (Orthologous group) and after all translate sequences removing partials. Annotation: Translated sequences were annotated using a combined strategy: KEGG, blast2go, and EggNOG. Quantification: The Kallisto software estimates the expression of assembled contigs. Genealogy inference: Gene selection used the orthologous group identifier of OrthoMCL as starting point. Sequences were aligned (MAFFT). HMMER search was performed to find genes selected in Amoebozoa available data. We performed genealogies using maximum-likelihood criteria with IQTREE.



4 General discussion and conclusions

In this dissertation, we began to generate tools to understand the ability of some testate amoebas to withstand adverse environmental conditions. Although the Arcellinida are a group of reasonable presence in the literature, it has only recently been the case that new large-scale sequencing data have emerged and few analyzes have attempted to understand the general functioning of these organisms metabolism and their influence on ecological processes (KANG et al., 2017; DUMACK et al., 2018; HOFSTATTER et al., 2016). The big data represented a huge point of entry into the functional understanding of the testate amoeba because it is capable of giving an overview of the story first and then addressing specific questions. The result of the present work was a description of the genomic scale data for *A.intermedia*, and the discovery of genes likely to influence the evolution of the testate-amoeba.

Chapter 2 of this work is the first assembled and annotated transcriptome of A.intermedia. Characterization of the functioning of *Arcella intermedia* was done by sequencing their transcriptome. Sampling of individual cells throughout the *A.intermedia* growth phases and also throughout the culture is of crucial importance in order to obtain a greater variety of cell moments. The result of this chapter was a description of the main metabolic pathways and the main functions developed by the *A.intermedia* cell. Many stories have arisen. However, the focus of this dissertation was (1) the evolution of genes related to anaerobic metabolism in tested amoeba lineages. The interest came from the fact that other microbial eukaryotes have the possibility that their organelles allow the development of both aerobic processes like oxidative phosphorylation and also the anaerobic metabolism depending on the situation in which the organism lives. As *Arcella* can live in many kinds of stressful conditions as for example, eutrophic environments, they are

likely to encounter anoxic or microaerophilic conditions frequently. Perhaps some Arcellinids could have some benefit if they had genes associated with anoxic survival. Here, we present the first evidence of anaerobic ATP generation pathways in some Arcelinids that are generally aerobic. By acquiring anaerobic energy metabolism pathways, they will be likely to survive in a wider range of habitats. In this work there is an indication of how some Arcelinids could survive under microaerophilic or eutrophic conditions.

In restricted niches, only organisms that evolved specific characteristics related to the environment must be able to survive. Our theory is that lateral transferences of genes have an enormous contribution in the conduction of the eukaryotic evolution. In the traditional view of evolution of eukaryotic genomes new features evolve mainly by a gene duplication followed by a series of mutations, functionally modifying the copy (neofunctionalization) (KOONIN; MAKAROVA; ARAVIND, 2001). However, this is a slow way of emergence of evolutionary novelties, since it requires an accumulation of mutations over several generations. On the other hand, LGT would allow an almost immediate acquisition of already optimized sequences, resulting in the new function in the recipient organism, allowing faster and drastic phenotypic changes. It has also been shown that transfer of a single or a few genes could have a large effect on the receptor phenotype at short evolutionary time scales (SCHÖNKNECHT; WEBER; LERCHER, 2014). Bacteria and archaea must have made great contributions in eukaryotic evolution through lateral gene transfers. This work will contribute with more evidence in the debate about lateral gene transfers in the evolution of microbial eukaryotes.

A.intermedia metabolism is consistent with the expected for eukaryotic organisms. However, in Chapter 3, we described anaerobic metabolism genes in *A.intermedia*. Like many other protist strains, some testate-amoebae probably have an alternate route for ATP generation under anaerobic or microaerophilic

conditions. Route known as extended glycolysis pathway. This pathway consists of genes related to anaerobic/microaerophylic metabolism (Acetyl-CoA synthase, PFO). Although we do not have all the complete sequences to be sure, we find mitochondrial target sequences in all *Difflugia* sequences. The use of codons for these genes is similar to that found for each species. Which is also a good indication of the genes belonging to the amoeba. The sequences of Arcellinida also appear grouped, indicating some degree of organismic relation. The genetics and location of the expression would still be interesting data for the subject. However, the work seems to show good indications of the expression of these genes in Arcelinids. In summary, our data indicate that Arcella, Difflugia and Cyclopyxis must be generating ATP by phosphorylation at the substrate level and synthesizing acetate in the mitochondria, perhaps also in the cytosol (with Acetyl-CoA synthase). Amoebozoa do not form a monophyletic group in any of the three genes. This is in agreement with works of other researchers, where the sequences of *Mastigamoeba balamuthi*, *Entamoeba histolytica* and *Acanthamoeba castelanii* do not form a monophyletic groups. Amoebozoa monophyly was also rejected by AU test.

The general conclusion of the work would be that the sequences of anaerobic metabolism present in Arcellinida strains should belong to the amoebae, and as expected by the literature of the area do not form a monophyletic grouping with other eukaryotes, having bacteria or archeas pervading the group. There are two scenarios that are most believed to originate from genes of anaerobic metabolism: these genes may have been acquired on the basis of eukaryotes, one or a few times and being lost independently among exclusive aerobic strains; Or they derive from multiple independent lateral transfers between organisms occupying the same type of environment (BARBERÀ et al., 2007). Our phylogenetic analysis then contributes evidence for an LGT origin of all three anaerobic energetic metabolism genes in Arcelinids. The results of this work are in agreement with previous

phylogenies of these genes (HUG; STECHMANN; ROGER, 2009; LEGER et al., 2016; STAIRS et al., 2014; NÝVLTOVÁ et al., 2015). Although present in the PFO and hydrogenase genes, it was not possible with our data to associate alphaproteobacteria directly with the entry of these genes into eukaryotes. Even after attempting to force such grouping, the AU test rejected this possibility. These evidences argue against the hypothesis of hydrogen (MARTIN; MÜLLER, 1998), where anaerobic metabolisms would have been acquired in the same event of mitochondria endosymbiosis. Although none of our analyzes made it possible to determine the bacterial sibling group of these anaerobic metabolism genes, they brought yet another independent group in which anaerobic metabolisms appeared, the Tubulinea, and further allow to infer that such acquisitions should not have happened once. Collecting more data on eukaryotes can improve the phylogenetic signal of these trees and perhaps clarify the origins of these proteins.

Bibliography

ADL, M.; GUPTA, V. S. Protists in soil ecology and forest nutrient cycling. *Canadian Journal of Forest Research*, NRC Research Press, v. 36, n. 7, p. 1805–1817, 2006. Citado na página 17.

AL-RUBEAI, M.; SINGH, R. P. Apoptosis in cell culture. *Current opinion in Biotechnology*, Elsevier, v. 9, n. 2, p. 152–156, 1998. Citado na página 20.

ALFIERI, C.; ZHANG, S.; BARFORD, D. Visualizing the complex functions and mechanisms of the anaphase promoting complex/cyclosome (apc/c). *Open biology*, Royal Society Journals, v. 7, n. 11, p. 170204, 2017. Citado na página 51.

ANANTHARAMAN, V.; IYER, L. M.; ARAVIND, L. Comparative genomics of protists: new insights into the evolution of eukaryotic signal transduction and gene regulation. *Annu. Rev. Microbiol.*, Annual Reviews, v. 61, p. 453–475, 2007. Citado na página 48.

ANDERSON, O. R. *Comparative protozoology: ecology, physiology, life history*. [S.l.]: Springer Science & Business Media, 2013. Citado na página 28.

ANDREY, T.; YURI, M. Morphology and biometry of arcella intermedia (deflandre, 1928) comb. nov.. from russia and a review of hemispheric species of the genus arcella (testcealobosea, arcellinida). *Protistology*, v. 4, n. 4, 2006. Citado na página 27.

ANDREY, T.; YURI, M. Morphology, biometry and ecology of arcella gibbosa penard 1890 (rhizopoda, testacealobosea). *Protistology*, v. 4, n. 3, 2006. Citado 2 vezes nas páginas 27 and 34.

BARBERÀ, M. J. et al. The diversity of mitochondrion-related organelles amongst eukaryotic microbes. In: *Origin of mitochondria and hydrogenosomes*. [S.l.]: Springer, 2007. p. 239–275. Citado na página 81.

BERTOLI, C.; SKOTHEIM, J. M.; BRUIN, R. A. D. Control of cell cycle transcription during g1 and s phases. *Nature reviews Molecular cell biology*, Nature Publishing Group, v. 14, n. 8, p. 518, 2013. Citado na página 50.

BEYENS, L.; MEISTERFELD, R. Protozoa: testate amoebae. In: *Tracking environmental change using lake sediments*. [S.l.]: Springer, 2002. p. 121–153. Citado na página 27.

BHATTACHARYA, D.; HELMCHEN, T.; MELKONIAN, M. Molecular evolutionary analyses of nuclear-encoded small subunit ribosomal rna identify an

independent rhizopod lineage containing the euglyphina and the chlorarachniophyta. *Journal of Eukaryotic Microbiology*, Wiley Online Library, v. 42, n. 1, p. 65–69, 1995. Citado na página 27.

BLANDENIER, Q. et al. Nad9/nad7 (mitochondrial nicotinamide adenine dinucleotide dehydrogenase gene)—a new "holy grail" phylogenetic and dna-barcoding marker for arcellinida (amoebozoa)? *European journal of protistology*, Elsevier, v. 58, p. 175–186, 2017. Citado na página 35.

BOEHM, M.; BONIFACINO, J. S. Adaptins: the final recount. *Molecular biology of the cell*, Am Soc Cell Biol, v. 12, n. 10, p. 2907–2920, 2001. Citado na página 49.

BOENIGK, J. et al. Concepts in protistology: species definitions and boundaries. *European Journal of Protistology*, Elsevier, v. 48, n. 2, p. 96–102, 2012. Citado na página 17.

BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics*, Oxford University Press, v. 30, n. 15, p. 2114–2120, 2014. Citado na página 59.

BOWERS, B.; KORN, E. D. The fine structure of acanthamoeba castellanii: I. the trophozoite. *The Journal of Cell Biology*, Rockefeller University Press, v. 39, n. 1, p. 95–111, 1968. Citado na página 45.

BOXMA, B. et al. An anaerobic mitochondrion that produces hydrogen. *Nature*, Nature Publishing Group, v. 434, n. 7029, p. 74, 2005. Citado 2 vezes nas páginas 77 and 78.

BROWN, M. W. et al. Phylogenomics demonstrates that breviate flagellates are related to opisthokonts and apusomonads. *Proceedings of the Royal Society of London B: Biological Sciences*, The Royal Society, v. 280, n. 1769, p. 20131755, 2013. Citado na página 27.

BROWN, T. *Genomes 2nd edition*. [S.l.]: BIOS Scientific Publishers, 2002. Citado 2 vezes nas páginas 23 and 24.

CASTILHO, L. et al. *Animal cell technology: from biopharmaceuticals to gene therapy*. [S.l.]: Garland Science, 2008. Citado 3 vezes nas páginas 18, 19, and 20.

CASTRESANA, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular biology and evolution*, Oxford University Press, v. 17, n. 4, p. 540–552, 2000. Citado na página 58.

CASTRESANA, J. Topological variation in single-gene phylogenetic trees. *Genome biology*, BioMed Central, v. 8, n. 6, p. 216, 2007. Citado 2 vezes nas páginas 26 and 76.

CATALANOTTI, C. et al. Fermentation metabolism and its evolution in algae. *Frontiers in plant science*, Frontiers Media SA, v. 4, 2013. Citado 6 vezes nas páginas 21, 46, 56, 75, 77, and 78.

CAVALIER-SMITH, T.; CHAD, E. Sarcomonad ribosomal rna sequences, rhizopod phylogeny, and the origin of euglyphid amoebae. *Archiv für Protistenkunde*, Elsevier, v. 147, n. 3-4, p. 227–236, 1997. Citado na página 27.

CLARKE, M. et al. Genome of acanthamoeba castellanii highlights extensive lateral gene transfer and early evolution of tyrosine kinase signaling. *Genome biology*, BioMed Central, v. 14, n. 2, p. R11, 2013. Citado na página 35.

CLAROS, M. G. Mitoprot, a macintosh application for studying mitochondrial proteins. *Bioinformatics*, Oxford University Press, v. 11, n. 4, p. 441–447, 1995. Citado na página 60.

CONESA, A.; GÖTZ, S. Blast2go: A comprehensive suite for functional analysis in plant genomics. *International journal of plant genomics*, Hindawi, v. 2008, 2008. Citado 2 vezes nas páginas 31 and 38.

CORLISS, J. O. Biodiversity and biocomplexity of the protists and an overview of their significant roles in maintenance of our biosphere. *Acta Protozoologica*, POLISH ACADEMY OF SCIENCES, v. 41, n. 3, p. 199–220, 2002. Citado na página 17.

CRISCUOLO, A.; GRIBALDO, S. Bmge (block mapping and gathering with entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC evolutionary biology*, BioMed Central, v. 10, n. 1, p. 210, 2010. Citado na página 60.

CROSS, F. R.; BUCHLER, N. E.; SKOTHEIM, J. M. Evolution of networks and sequences in eukaryotic cell cycle control. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, The Royal Society, v. 366, n. 1584, p. 3532–3544, 2011. Citado na página 50.

DEFLANDRE, G. Le genre arcella ehrenberg. *Archiv für Protistenkunde*, v. 64, n. 1, p. 152–287, 1928. Citado 2 vezes nas páginas 27 and 34.

DEPONTE, M. Programmed cell death in protists. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, Elsevier, v. 1783, n. 7, p. 1396–1405, 2008. Citado na página 51.

DIJKEN, J. P. van; SCHEFFERS, W. A. Redox balances in the metabolism of sugars by yeasts. *FEMS microbiology letters*, Elsevier, v. 32, n. 3-4, p. 199–224, 1986. Citado na página 77.

DUMACK, K. et al. Reinvestigation of phryganella paradoxa (arcellinida, amoebozoa) penard 1902. *Journal of Eukaryotic Microbiology*, Wiley Online Library, 2018. Citado na página 79.

EICHINGER, L. et al. The genome of the social amoeba dictyostelium discoideum. *Nature*, Nature Publishing Group, v. 435, n. 7038, p. 43, 2005. Citado na página 35.

ELLIS, J. E. et al. Electron transport components of the parasitic protozoon giardia lamblia. *FEBS letters*, Elsevier, v. 325, n. 3, p. 196–200, 1993. Citado na página 21.

EMANUELSSON, O. et al. Locating proteins in the cell using targetp, signalp and related tools. *Nature protocols*, Nature Publishing Group, v. 2, n. 4, p. 953, 2007. Citado na página 60.

ESCOBAR, J. et al. Ecology of testate amoebae (thecamoebians) in subtropical florida lakes. *Journal of Paleolimnology*, Springer, v. 40, n. 2, p. 715–731, 2008. Citado na página 58.

FENCHEL, T. Anaerobic eukaryotes. In: *Anoxia*. [S.l.]: Springer, 2012. p. 3–16. Citado na página 56.

FENCHEL, T.; FINLAY, B. *Ecology and Evolution in Anoxic Worlds*. Oxford University Press, 1995. (Oxford series in ecology and evolution). ISBN 9780198548379. Disponível em: ⟨https://books.google.com.br/books?id=7nfwAAAAMAAJ⟩. Citado na página 56.

FERES, J. C. et al. Morphological and morphometric description of a novel shelled amoeba arcella gandalfi sp. nov.(amoebozoa: Arcellinida) from brazilian continental waters. *Acta Protozoologica*, Portal Czasopism Naukowych Ejournals. eu, v. 2016, n. 4, p. 221–229, 2016. Citado 2 vezes nas páginas 27 and 34.

FONTANETO, D. *Biogeography of microscopic organisms: is everything small everywhere?* [S.l.]: Cambridge University Press, 2011. Citado 2 vezes nas páginas 34 and 74.

FRESHNEY, R. I. *Culture of animal cells, a manual of basic technique*. [S.l.]: John Wiley and Sons, inc, 1994. Citado na página 19.

FRITZ-LAYLIN, L. K. et al. The genome of naegleria gruberi illuminates early eukaryotic versatility. *Cell*, Elsevier, v. 140, n. 5, p. 631–642, 2010. Citado na página 57.

FUERST, J. A. Beyond prokaryotes and eukaryotes: planctomycetes and cell organization. *Nat. Educ*, v. 3, n. 9, p. 44, 2010. Citado na página 22.

GEISEN, S. et al. Soil protists: a fertile frontier in soil biology research. *FEMS microbiology reviews*, 2018. Citado na página 24.

GINGER, M. L. et al. Intermediary metabolism in protists: a sequence-based view of facultative anaerobic metabolism in evolutionarily diverse eukaryotes. *Protist*, Elsevier, v. 161, n. 5, p. 642–671, 2010. Citado 2 vezes nas páginas 21 and 57.

GLEISNER, M. et al. Epsin n-terminal homology domain (enth) activity as a function of membrane tension. *Journal of Biological Chemistry*, ASBMB, p. jbc–M116, 2016. Citado na página 49.

GOUY, M.; GUINDON, S.; GASCUEL, O. Seaview version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular biology and evolution*, Oxford University Press, v. 27, n. 2, p. 221–224, 2009. Citado 2 vezes nas páginas 31 and 60.

GRABHERR, M. G. et al. Full-length transcriptome assembly from rna-seq data without a reference genome. *Nature biotechnology*, Nature Publishing Group, v. 29, n. 7, p. 644, 2011. Citado na página 59.

GRANT, B. D.; DONALDSON, J. G. Pathways and mechanisms of endocytic recycling. *Nature reviews Molecular cell biology*, Nature Publishing Group, v. 10, n. 9, p. 597, 2009. Citado na página 50.

GRANT, J. R.; KATZ, L. A. Building a phylogenomic pipeline for the eukaryotic tree of life-addressing deep phylogenies with genome-scale data. *PLoS currents*, Public Library of Science, v. 6, 2014. Citado 2 vezes nas páginas 31 and 59.

GRANT, J. R. et al. Gene discovery from a pilot study of the transcriptomes from three diverse microbial eukaryotes: Corallomyxa tenera, chilodonella uncinata, and subulatomonas tetraspora. *Protist Genomics*, Versita, v. 1, p. 3–18, 2012. Citado 2 vezes nas páginas 24 and 25.

HAMPL, V.; STAIRS, C. W.; ROGER, A. J. The tangled past of eukaryotic enzymes involved in anaerobic metabolism. *Mobile genetic elements*, Taylor & Francis, v. 1, n. 1, p. 71–74, 2011. Citado na página 57.

HAUSER, H.; WAGNER, R. *Mammalian cell biotechnology in protein production*. [S.l.]: de Gruyter, 1997. Citado na página 20.

HECKMAN, C. W. *The Pantanal of Poconé: Biota and ecology in the northern section of the world's largest pristine wetland*. [S.l.]: Springer Science & Business Media, 2013. v. 77. Citado na página 74.

HEGNER, R. Heredity, variation, and the appearance of diversities during the vegetative reproduction of arcella dentata. *Genetics*, Genetics Society of America, v. 4, n. 2, p. 95, 1919. Citado na página 34.

HEGNER, R. W. The relations between nuclear number, chromatin mass, cytoplasmic mass, and shell characteristics in four pieces of the genus arcella. *Journal of Experimental Zoology*, Wiley Online Library, v. 30, n. 1, p. 1–95, 1920. Citado na página 34.

HILLMANN, F. et al. Multiple roots of fruiting body formation in amoebozoa. *Genome biology and evolution*, Oxford University Press, v. 10, n. 2, p. 591–606, 2018. Citado na página 25.

HOFFENBERG, S. et al. A novel membrane-anchored rab5 interacting protein required for homotypic endosome fusion. *Journal of Biological Chemistry*, ASBMB, v. 275, n. 32, p. 24661–24669, 2000. Citado na página 49.

HOFSTATTER, P. G. et al. Evolution of bacterial recombinase a (reca) in eukaryotes explained by addition of genomic data of key microbial lineages. *Proc. R. Soc. B*, The Royal Society, v. 283, n. 1840, p. 20161453, 2016. Citado 3 vezes nas páginas 34, 35, and 79.

HUG, L. A.; STECHMANN, A.; ROGER, A. J. Phylogenetic distributions and histories of proteins involved in anaerobic pyruvate metabolism in eukaryotes. *Molecular biology and evolution*, Oxford University Press, v. 27, n. 2, p. 311–324, 2009. Citado 2 vezes nas páginas 75 and 82.

HUTAGALUNG, A. H.; NOVICK, P. J. Role of rab gtpases in membrane traffic and cell physiology. *Physiological reviews*, American Physiological Society Bethesda, MD, v. 91, n. 1, p. 119–149, 2011. Citado na página 50.

JAN, P. The twilight of sarcodina: a molecular perspective on the polyphyletic origin of amoeboid protists. *Protistology*, v. 5, n. 4, 2008. Citado na página 27.

JENSEN, L. J. et al. eggnog: automated construction and annotation of orthologous groups of genes. *Nucleic acids research*, Oxford University Press, v. 36, n. suppl_1, p. D250–D254, 2007. Citado na página 38.

KAESSMANN, H. Origins, evolution and phenotypic impact of new genes. *Genome research*, Cold Spring Harbor Lab, p. gr–101386, 2010. Citado na página 23.

KANEHISA, M.; SATO, Y.; MORISHIMA, K. Blastkoala and ghostkoala: Kegg tools for functional characterization of genome and metagenome sequences. *Journal of molecular biology*, Elsevier, v. 428, n. 4, p. 726–731, 2016. Citado na página 38.

KANG, S. et al. Between a pod and a hard test: the deep evolution of amoebae. *Molecular biology and evolution*, Oxford University Press, v. 34, n. 9, p. 2258–2270, 2017. Citado 5 vezes nas páginas 25, 27, 36, 59, and 79.

KATOH, K.; ASIMENOS, G.; TOH, H. Multiple alignment of dna sequences with mafft. In: *Bioinformatics for DNA sequence analysis*. [S.l.]: Springer, 2009. p. 39–64. Citado 3 vezes nas páginas 31, 58, and 60.

KEELING, P. J.; CAMPO, J. del. Marine protists are not just big bacteria. *Current Biology*, Elsevier, v. 27, n. 11, p. R541–R549, 2017. Citado na página 17.

KIM, S. et al. De novo transcriptome analysis of an arctic microalga, chlamydomonas sp. *Genes & Genomics*, Springer, v. 35, n. 2, p. 215–223, 2013. Citado na página 25.

KOONIN, E. V.; MAKAROVA, K. S.; ARAVIND, L. Horizontal gene transfer in prokaryotes: quantification and classification. *Annual Reviews in Microbiology*, Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA, v. 55, n. 1, p. 709–742, 2001. Citado 2 vezes nas páginas 24 and 80.

KOSAKYAN, A. et al. Current and future perspectives on the systematics, taxonomy and nomenclature of testate amoebae. *European journal of protistology*, Elsevier, v. 55, p. 105–117, 2016. Citado na página 28.

KOURANTI, I. et al. Rab35 regulates an endocytic recycling pathway essential for the terminal steps of cytokinesis. *Current biology*, Elsevier, v. 16, n. 17, p. 1719–1725, 2006. Citado na página 50.

KREBS, H.; HENSELEIT, K. et al. Formation of urea in the animal body. *Hoppe-Seyler's Zeitschrift fur physiologische Chemie*, v. 210, p. 33–66, 1932. Citado na página 46.

KREBS, H. A. The discovery of the ornithine cycle of urea synthesis. *Biochemical Education*, Wiley Online Library, v. 1, n. 2, p. 19–23, 1973. Citado na página 46.

KURLAND, C.; ANDERSSON, S. Origin and evolution of the mitochondrial proteome. *Microbiology and Molecular Biology Reviews*, Am Soc Microbiol, v. 64, n. 4, p. 786–820, 2000. Citado na página 22.

LAHR, D. J. et al. Phylogenomics and ancestral morphological reconstruction of testate amoebae demonstrate high diversity of microbial eukaryotes in the neoproterozoic. *Current biology*. Citado na página 35.

LANSAC-TÔHA, F. et al. Structure of the testate amoebae community in different habitats in a neotropical floodplain. *Brazilian Journal of Biology*, SciELO Brasil, v. 74, n. 1, p. 181–190, 2014. Citado na página 28.

LEGER, M. M. et al. Novel hydrogenosomes in the microaerophilic jakobid stygiella incarcerata. *Molecular biology and evolution*, Oxford University Press, v. 33, n. 9, p. 2318–2336, 2016. Citado 2 vezes nas páginas 75 and 82.

LI, L.; STOECKERT, C. J.; ROOS, D. S. Orthomcl: identification of ortholog groups for eukaryotic genomes. *Genome research*, Cold Spring Harbor Lab, v. 13, n. 9, p. 2178–2189, 2003. Citado na página 58.

LOFTUS, B. et al. The genome of the protist parasite entamoeba histolytica. *Nature*, Nature Publishing Group, v. 433, n. 7028, p. 865, 2005. Citado na página 35.

MACUMBER, A. L. et al. Autoecological approaches to resolve subjective taxonomic divisions within arcellacea. *Protist*, Elsevier, v. 165, n. 3, p. 305–316, 2014. Citado na página 28.

MADONI, P. Protozoa in wastewater treatment processes: a minireview. *Italian Journal of Zoology*, Taylor & Francis, v. 78, n. 1, p. 3–11, 2011. Citado na página 57.

MAIER, R. M.; PEPPER, I. L.; GERBA, C. P. *Environmental microbiology*. [S.l.]: Academic press, 2009. v. 397. Citado na página 17.

MANNING, B. D.; CANTLEY, L. C. Akt/pkb signaling: navigating downstream. *Cell*, Elsevier, v. 129, n. 7, p. 1261–1274, 2007. Citado na página 47.

MARCHADIER, E. et al. Evolution of the calcium-based intracellular signaling system. *Genome biology and evolution*, Oxford University Press, v. 8, n. 7, p. 2118–2132, 2016. Citado na página 48.

MARTIN, W.; MÜLLER, M. The hydrogen hypothesis for the first eukaryote. *Nature*, Nature Publishing Group, v. 392, n. 6671, p. 37, 1998. Citado 2 vezes nas páginas 58 and 82.

MASAI, H. et al. Eukaryotic chromosome dna replication: where, when, and how? *Annual review of biochemistry*, v. 79, 2010. Citado na página 50.

MEDINGER, R. et al. Diversity in a hidden world: potential and limitation of next-generation sequencing for surveys of molecular diversity of eukaryotic microorganisms. *Molecular ecology*, Wiley Online Library, v. 19, n. s1, p. 32–40, 2010. Citado na página 17.

MEDIOLI, F.; SCOTT, D. B. Lacustrine thecamoebians (mainly arcellaceans) as potential tools for palaeolimnological interpretations. *Palaeogeography, Palaeoclimatology, Palaeoecology*, Elsevier, v. 62, n. 1-4, p. 361–386, 1988. Citado na página 74.

MIGNOT, J.-P.; RAIKOV, I. B. Evidence for meiosis in the testate amoeba arcella. *The Journal of protozoology*, Wiley Online Library, v. 39, n. 2, p. 287–289, 1992. Citado 2 vezes nas páginas 34 and 35.

MIHAYLOVA, D.; PROKOPOV, T.; MIHALKOV, N. Ecologia balkanica. Citado na página 28.

MORSE, D.; DAOUST, P.; BENRIBAGUE, S. A transcriptome-based perspective of cell cycle regulation in dinoflagellates. *Protist*, Elsevier, v. 167, n. 6, p. 610–621, 2016. Citado 2 vezes nas páginas 50 and 51.

MÜLLER, M. Energy metabolism: Part i: Anaerobic protozoa. In: *Molecular Medical Parasitology*. [S.l.]: Elsevier, 2003. p. 125–139. Citado 2 vezes nas páginas 77 and 78.

MÜLLER, M. et al. Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *Microbiology and Molecular Biology Reviews*, Am Soc Microbiol, v. 76, n. 2, p. 444–495, 2012. Citado 3 vezes nas páginas 22, 57, and 77.

MUS, F. et al. Anaerobic acclimation in chlamydomonas reinhardtii anoxic gene expression, hydrogenase induction, and metabolic pathways. *Journal of Biological Chemistry*, ASBMB, v. 282, n. 35, p. 25475–25486, 2007. Citado na página 45.

NASSER, N. A. et al. Lacustrine arcellinina (testate amoebae) as bioindicators of arsenic contamination. *Microbial ecology*, Springer, v. 72, n. 1, p. 130–149, 2016. Citado na página 28.

NGUYEN, L.-T. et al. Iq-tree: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular biology and evolution*, Oxford University Press, v. 32, n. 1, p. 268–274, 2014. Citado na página 61.

NIKOLAEV, S. I. et al. The testate lobose amoebae (order arcellinida kent, 1880) finally find their home within amoebozoa. *Protist*, Elsevier, v. 156, n. 2, p. 191–202, 2005. Citado na página 27.

NÝVLTOVÁ, E. et al. Lateral gene transfer and gene duplication played a key role in the evolution of mastigamoeba balamuthi hydrogenosomes. *Molecular biology and evolution*, Oxford University Press, v. 32, n. 4, p. 1039–1055, 2015. Citado 8 vezes nas páginas 21, 22, 56, 57, 67, 75, 77, and 82.

NÝVLTOVÁ, E. et al. Nif-type iron-sulfur cluster assembly system is duplicated and distributed in the mitochondria and cytosol of mastigamoeba balamuthi. *Proceedings of the National Academy of Sciences*, National Acad Sciences, v. 110, n. 18, p. 7371–7376, 2013. Citado na página 35.

OLDHAM, W. M.; HAMM, H. E. Heterotrimeric g protein activation by g-protein-coupled receptors. *Nature reviews Molecular cell biology*, Nature Publishing Group, v. 9, n. 1, p. 60, 2008. Citado na página 48.

OLSON, L. K. et al. *Genomics and evolution of microbial eukaryotes*. [S.l.]: Oxford University Press on Demand, 2006. Citado na página 24.

OPPERDOES, F. R.; JONCKHEERE, J. F. D.; TIELENS, A. G. Naegleria gruberi metabolism. *International journal for parasitology*, Elsevier, v. 41, n. 9, p. 915–924, 2011. Citado 3 vezes nas páginas 45, 46, and 77.

PATNAIK, B. B. et al. Sequencing, de novo assembly, and annotation of the transcriptome of the endangered freshwater pearl bivalve, cristaria plicata, provides novel insights into functional genes and marker discovery. *PloS one*, Public Library of Science, v. 11, n. 2, p. e0148622, 2016. Citado na página 25.

PAWLOWSKI, J. et al. Cbol protist working group: barcoding eukaryotic richness beyond the animal, plant, and fungal kingdoms. *PLoS biology*, Public Library of Science, v. 10, n. 11, p. e1001419, 2012. Citado na página 17.

PAWLOWSKI, J. et al. Environmental monitoring through protist next-generation sequencing metabarcoding: assessing the impact of fish farming on benthic foraminifera communities. *Molecular Ecology Resources*, Wiley Online Library, v. 14, n. 6, p. 1129–1140, 2014. Citado na página 17.

PAWLOWSKI, L.; DUDZINSKA, M.; PAWLOWSKI, A. *Environmental Engineering III*. CRC Press, 2010. ISBN 9780203846667. Disponível em: ⟨https://books.google.com.br/books?id=_NXKBQAAQBAJ⟩. Citado na página 28.

PAYNE, S. H.; LOOMIS, W. F. Retention and loss of amino acid biosynthetic pathways based on analysis of whole-genome sequences. *Eukaryotic cell*, Am Soc Microbiol, v. 5, n. 2, p. 272–276, 2006. Citado na página 47.

PCHELIN, I. M. Testate amoeba arcella vulgaris (amoebozoa, arcellidae) is able to survive without the shell and construct a new one. *Protistology*, v. 6, n. 4, p. 251–257, 2010. Citado na página 27.

PFEIFFER, T.; SCHUSTER, S.; BONHOEFFER, S. Cooperation and competition in the evolution of atp-producing pathways. *Science*, American Association for the Advancement of Science, v. 292, n. 5516, p. 504–507, 2001. Citado na página 77.

PICELLI, S. et al. Full-length rna-seq from single cells using smart-seq2. *Nature protocols*, Nature Publishing Group, v. 9, n. 1, p. 171, 2014. Citado 2 vezes nas páginas 30 and 37.

PINEDA, E. et al. Pyruvate: ferredoxin oxidoreductase and bifunctional aldehyde-alcohol dehydrogenase are essential for energy metabolism under oxidative stress in entamoeba histolytica. *The FEBS journal*, Wiley Online Library, v. 277, n. 16, p. 3382–3395, 2010. Citado na página 77.

PORFÍRIO-SOUSA, A. L.; RIBEIRO, G. M.; LAHR, D. J. Morphometric and genetic analysis of arcella intermedia and arcella intermedia laevis (amoebozoa, arcellinida) illuminate phenotypic plasticity in microbial eukaryotes. *European journal of protistology*, Elsevier, v. 58, p. 187–194, 2017. Citado 2 vezes nas páginas 34 and 36.

QIU, H. et al. Adaptation through horizontal gene transfer in the cryptoendolithic red alga galdieria phlegrea. *Current Biology*, Elsevier, v. 23, n. 19, p. R865–R866, 2013. Citado na página 25.

RAGAN, M. *A biochemical phylogeny of the protists*. [S.l.]: Elsevier, 2012. Citado na página 46.

RAO, H. et al. Degradation of a cohesin subunit by the n-end rule pathway is essential for chromosome stability. *Nature*, Nature Publishing Group, v. 410, n. 6831, p. 955, 2001. Citado na página 51.

RECZUGA, M. K. et al. Arcella peruviana sp. nov.(amoebozoa: Arcellinida, arcellidae), a new species from a tropical peatland in amazonia. *European journal of protistology*, Elsevier, v. 51, n. 5, p. 437–449, 2015. Citado na página 34.

REINHARD, C. T. et al. Earth's oxygen cycle and the evolution of animal life. *Proceedings of the National Academy of Sciences*, National Academy of Sciences, v. 113, n. 32, p. 8933–8938, 2016. ISSN 0027-8424. Disponível em: ⟨http://www.pnas.org/content/113/32/8933⟩. Citado 2 vezes nas páginas 21 and 56.

REYNOLDS, B. D. Inheritance of double characteristics in arcella polypora penard. *Genetics*, Genetics Society of America, v. 8, n. 5, p. 477, 1923. Citado na página 34.

RIDE, W. et al. *International code of zoological nomenclature*. [S.l.]: International Trust for Zoological Nomenclature, 1999. Citado na página 27.

RISMANI-YAZDI, H. et al. Transcriptome sequencing and annotation of the microalgae dunaliella tertiolecta: pathway description and gene discovery for production of next-generation biofuels. *BMC genomics*, BioMed Central, v. 12, n. 1, p. 148, 2011. Citado na página 25.

RODRÍGUEZ, M. A. et al. Cloning and characterization of the entamoeba histolytica pyruvate: ferredoxin oxidoreductase gene. *Molecular and biochemical parasitology*, Elsevier, v. 78, n. 1-2, p. 273–277, 1996. Citado na página 21.

ROE, H. M.; PATTERSON, R. T. Arcellacea (testate amoebae) as bio-indicators of road salt contamination in lakes. *Microbial ecology*, Springer, v. 68, n. 2, p. 299–313, 2014. Citado na página 28.

ROE, H. M.; PATTERSON, R. T.; SWINDLES, G. T. Controls on the contemporary distribution of lake thecamoebians (testate amoebae) within the greater toronto area and their potential as water quality indicators. *Journal of Paleolimnology*, Springer, v. 43, n. 4, p. 955–975, 2010. Citado 2 vezes nas páginas 28 and 57.

RUSSELL, B. et al. The directiveness of organic activities. *The directiveness of organic activities.*, CambridgeUniversity Press, 1944. Citado na página 74.

SAGAN, L. On the origin of mitosing cells. *Journal of theoretical biology*, Elsevier, v. 14, n. 3, p. 225–IN6, 1967. Citado na página 22.

SANCHEZ, L. B.; MÜLLER, M. Purification and characterization of the acetate forming enzyme, acetyl-coa synthetase (adp-forming) from the amitochondriate protist, giardia lamblia. *FEBS letters*, Wiley Online Library, v. 378, n. 3, p. 240–244, 1996. Citado na página 77.

SAXTON, R. A.; SABATINI, D. M. mtor signaling in growth, metabolism, and disease. *Cell*, Elsevier, v. 168, n. 6, p. 960–976, 2017. Citado na página 48.

SCHAAP, P. et al. The physarum polycephalum genome reveals extensive use of prokaryotic two-component and metazoan-type tyrosine kinase signaling. *Genome biology and evolution*, Oxford University Press, v. 8, n. 1, p. 109–125, 2015. Citado na página 35.

SCHIMPER, A. F. W. On the development of the chlorophyll grains and color bodies. *bot time*, v. 41, p. 105–112, 1883. Citado na página 22.

SCHÖNKNECHT, G.; WEBER, A. P.; LERCHER, M. J. Horizontal gene acquisitions by eukaryotes as drivers of adaptive evolution. *Bioessays*, Wiley Online Library, v. 36, n. 1, p. 9–20, 2014. Citado 2 vezes nas páginas 24 and 80.

SCOTT, D. B.; MEDIOLI, F. S.; SCHAFER, C. T. *Monitoring in coastal environments using foraminifera and thecamoebian indicators*. [S.l.]: Cambridge University Press, 2007. Citado na página 28.

SHAW, R. J. Lkb1 and amp-activated protein kinase control of mtor signalling and growth. *Acta physiologica*, Wiley Online Library, v. 196, n. 1, p. 65–80, 2009. Citado na página 47.

SIKDAR, S. *Fundamentals and Applications of Bioremediation: Principles*. [S.l.]: Routledge, 2017. v. 1. Citado 2 vezes nas páginas 21 and 56.

SILVA, C. d. L. et al. Evaluation of sediment contamination by trace elements and the zooplankton community analysis in area affected by gold exploration in southeast (se) of the iron quadrangle, alto rio doce,(mg) brazil. *Acta Limnologica Brasiliensia*, SciELO Brasil, v. 25, n. 2, p. 150–157, 2013. Citado na página 28.

SINGH, R. et al. Improved annotation with de novo transcriptome assembly in four social amoeba species. *BMC genomics*, BioMed Central, v. 18, n. 1, p. 120, 2017. Citado na página 25.

ŠLAPETA, J.; MOREIRA, D.; LÓPEZ-GARCÍA, P. The extent of protist diversity: insights from molecular ecology of freshwater eukaryotes. *Proceedings of the Royal Society of London B: Biological Sciences*, The Royal Society, v. 272, n. 1576, p. 2073–2081, 2005. Citado na página 17.

SONG, S. et al. Small gtpases: structure, biological function and its interaction with nanoparticles. *Asian Journal of Pharmaceutical Sciences*, Elsevier, 2018. Citado na página 48.

STAIRS, C. W. et al. A suf fe-s cluster biogenesis system in the mitochondrion-related organelles of the anaerobic protist pygsuia. *Current Biology*, Elsevier, v. 24, n. 11, p. 1176–1186, 2014. Citado na página 82.

SUN, Y. et al. Dna damage-induced acetylation of lysine 3016 of atm activates atm kinase activity. *Molecular and cellular biology*, Am Soc Microbiol, v. 27, n. 24, p. 8502–8509, 2007. Citado na página 51.

SZELECZ, I. et al. Can soil testate amoebae be used for estimating the time since death? a field experiment in a deciduous forest. *Forensic science international*, Elsevier, v. 236, p. 90–98, 2014. Citado na página 17.

TICE, A. K. et al. Expansion of the molecular and morphological diversity of acanthamoebidae (centramoebida, amoebozoa) and identification of a novel life cycle type within the group. *Biology direct*, BioMed Central, v. 11, n. 1, p. 69, 2016. Citado na página 59.

TODOROV, M.; GOLEMANSKY, V. Morphology, biometry and ecology of arcella excavata cunningham, 1919 (rhizopoda: Arcellinida). *Acta protozoologica*, POLISH ACADEMY OF SCIENCES, v. 42, n. 2, p. 105–112, 2003. Citado na página 34.

TSAOUSIS, A. D. et al. The biochemical adaptations of mitochondrion-related organelles of parasitic and free-living microbial eukaryotes to low oxygen environments. In: *Anoxia*. [S.l.]: Springer, 2012. p. 51–81. Citado na página 22.

WEBER, C. et al. Extensive transcriptome analysis correlates the plasticity of entamoeba histolytica pathogenesis to rapid phenotype changes depending on the environment. *Scientific reports*, Nature Publishing Group, v. 6, p. 35852, 2016. Citado na página 25.

WILKINSON, D. M.; CREEVY, A. L.; VALENTINE, J. The past, present and future of soil protist ecology. *Acta Protozoologica*, Jagiellonian University-Jagiellonian University Press, v. 51, n. 3, p. 189, 2012. Citado na página 17.

WILLIAMS, K.; LOWE, P.; LEADLAY, P. Purification and characterization of pyruvate: ferredoxin oxidoreductase from the anaerobic protozoon trichomonas vaginalis. *Biochemical Journal*, Portland Press Limited, v. 246, n. 2, p. 529–536, 1987. Citado na página 21.

YOSHIDA, Y. et al. De novo assembly and comparative transcriptome analysis of euglena gracilis in response to anaerobic conditions. *BMC genomics*, BioMed Central, v. 17, n. 1, p. 182, 2016. Citado na página 25.

ZHANG, M. et al. Rab7: roles in membrane trafficking and disease. *Bioscience reports*, Portland Press Limited, v. 29, n. 3, p. 193–209, 2009. Citado na página 50.

ZHU, R. et al. De novo annotation of the immune-enriched transcriptome provides insights into immune system genes of chinese sturgeon (acipenser sinensis). *Fish & shellfish immunology*, Elsevier, v. 55, p. 699–716, 2016. Citado na página 25.