

Título em Português:	PVT: uma ferramenta de visualização computacional para identificação de Amebas de Vida Livre
Título em Inglês:	pvt: a computational visualization tool for identification of free-living amoebas
Autor:	Matheus Issa
Instituição:	Universidade de São Paulo
Unidade:	Instituto de Física de São Carlos
Orientador:	Otávio Henrique Thiemann
Área de Pesquisa / SubÁrea:	Biologia e Fisiologia dos Microorganismos
Agência Financiadora:	FAPESP - Fundação de Amparo à Pesquisa do Estado de São Paulo

PVT: UMA FERRAMENTA DE VISUALIZAÇÃO COMPUTACIONAL PARA IDENTIFICAÇÃO DE AMEBAS DE VIDA LIVRE

Matheus Issa

Natália Karla Bellini, Douglas Cedrim Oliveira

Otavio Henrique Thiemann

Instituto de Física de São Carlos, Universidade de São Paulo

matheus.issa@usp.br

Objetivos

O presente trabalho tem, por objetivo, construir, de maneira sistemática e ordenada, um banco de dados robusto contendo todas as características de Amebas de Vida Livre (AVLs) para cada uma das espécies descritas no guia de classificação por chaves taxonômicas “A New Key to Freshwater and Soil Gymnamoebae” (1988), de Frederick C. Page, visando empregar uma refinada análise para elencar as informações essenciais para a classificação e distinção desses microrganismos [1]. Com isso, possibilita-se o desenvolvimento de uma ferramenta computacional de visualização, denominada *Page’s Visualization Tool* (PVT), com o intuito de elaborar um mecanismo capaz de aperfeiçoar os estudos morfológicos de AVLs atuais, os quais se baseiam na pesquisa e leitura manual do guia de classificação e, portanto, não são otimizados.

Métodos e Procedimentos

A elaboração da ferramenta computacional PVT se iniciou a partir da construção de um banco de dados organizando, sistematicamente, as descrições morfológicas para cada uma das espécies abordadas nas chaves de classificação do Page (1988) e as reunindo em tabelas. A tabulação dos dados foi feita em uma planilha dispondo todas as espécies descritas no guia nas linhas da

tabela, separadas por classe, e classificadas dentro de cada categoria explicitada pelo autor ao longo das chaves, arranjadas nas colunas da mesma. Foram inseridos os parâmetros específicos para as classes, ordens, famílias, gêneros e, por fim, espécies, de forma a abranger o máximo de informação possível para posterior análise refinada.

Uma vez finalizado o banco de dados, a construção do software em si se tornou possível. Tendo em vista que o banco de dados dispõe de dados multivariados, já que cada espécie se relaciona com cada uma das morfologias/características existentes e, portanto, trata-se de uma matriz com geometria de alta dimensão, optou-se por utilizar a técnica de visualização multidimensional de coordenadas paralelas para realizar a análise dos dados de forma efetiva [2-3].

Assim, a partir do arquivo .csv do banco de dados, utilizou-se a biblioteca “d3” na linguagem JavaScript, associada a recursos de desenvolvimento web, para gerar um programa rápido e interativo em coordenadas paralelas, em que as espécies das amebas se localizam nas linhas e cada uma das características morfológicas nas colunas.

Para o desenvolvimento da ferramenta PVT, o trabalho contou com a colaboração da Dra. Natália Karla Bellini, do Instituto Butantan, e do Prof. Dr. Douglas Cedrim Oliveira, do Instituto Federal Goiano (IF Goiano).

Resultados

A tabulação das informações contidas no guia de classificação Page gerou um banco de dados formado por 225 espécies distintas de Amebas de Vida Livre, associadas a 484 categorias distintas de morfologias e características gerais para tais AVLs. Elas foram distribuídas em informações dos estágios de vida, do corpo de frutificação das espécies que o possuem, das temperaturas de crescimento, dos meios de cultura utilizados para a obtenção das informações explicitadas, da descrição do estilo de vida/patogenicidade e da distribuição geográfica das espécies.

Uma cuidadosa investigação do banco de dados permitiu a constatação de que as informações acerca dos três estágios de vida das AVLs são fundamentais para a efetiva distinção morfológica entre as espécies, sendo as do estágio de trofozoíto as mais gerais e completas, especialmente na ausência de características dos estágios encistado e flagelado ou presença de poucas delas. Os dados relacionados aos cistos são essenciais para a diferenciação de espécies quando há uma plena descrição dos mesmos, ao passo que os associados às formas flageladas se tornam cruciais quando os de cistos não são totalmente suficientes. Em contrapartida, informações complementares que descrevem meios de cultura, dispersão geográfica, estilo de vida e temperatura de crescimento não apresentam alta significância, acrescentando especificações únicas para poucas espécies. As características das espécies que produzem corpo de frutificação, por sua vez, têm sua devida relevância por diferenciar a ordem Acrasida das demais.

Em posse dessas informações e do banco de dados, a aplicação das técnicas computacionais para uma análise preliminar permitiu a construção de uma interface interativa para o software PVT, de modo a possibilitar que o usuário insira manual e rapidamente as características observadas por microscopia em suas amostras e consiga obter, no melhor dos casos, a espécie estudada ou pelo menos o gênero a qual ela pertence.

Conclusões

A organização sistemática das informações contidas no guia Page possibilitou a construção

de um banco de dados robusto e volumoso contendo todas as descrições trazidas pelo autor em suas chaves de classificação. A análise refinada empregada permitiu averiguar categorias efetivas e relevantes na distinção entre as espécies, em especial aquelas associadas aos estágios de vida e, também, as características descritivas dos corpos de frutificação para as espécies que os produzem. As demais apresentaram impacto apenas para classificações pontuais. O estudo apurado do banco de dados serviu de alicerce para o desenvolvimento inovador do software PVT com uma interface interativa e eficiente que, embora já tenha mostrado eficácia no seu propósito, atualizações das informações são necessárias, especialmente para a adição de dados tanto morfológicos quanto moleculares recentes ao banco de dados. Com isso, possibilita-se o desenvolvimento de uma inovadora técnica otimizada, efetiva e qualificada para aprimorar os estudos de Amebas de Vida Livre de maneira geral, relacionando-se a abordagem morfológica e molecular para as análises. Dessa forma, acredita-se que o conhecimento acerca das AVLs poderá ser ampliado, suprimindo a escassez das pesquisas no tema e reduzindo os impactos socioambientais causados por essa problemática. Além disso, tal contribuição poderá auxiliar não só os trabalhos sobre essa temática como também poderá servir de modelo para a sistematização de características para outros organismos, sejam eles procariotos ou eucariotos.

Referências Bibliográficas

- 1 PAGE, F.C. **A new key to freshwater and soil Gymnamoebae**: with instructions for culture. Ambleside: Freshwater Biological Association, 1988. (Culture collection of algae and protozoa). ISBN 9781871105025.
- 2 HEINRICH, J.; WEISKOPF, D. State of the art of parallel coordinates. **Eurographics** - state of the art reports. 2013. p. 95-116,. Disponível em: https://journals.de/files/heinrich_state_2013.pdf. Acesso em: 05 set. 2022.
- 3 MUNZNER, T.. **Visualization analysis and design**. Boca Raton: CRC Press, 2014.

PVT: A COMPUTATIONAL VISUALIZATION TOOL FOR IDENTIFICATION OF FREE-LIVING AMOEBAS

Matheus Issa

Natália Karla Bellini, Douglas Cedrim Oliveira

Otavio Henrique Thiemann

São Carlos Institute of Physics, University of São Paulo

matheus.issa@usp.br

Objectives

The present work aims to build, systematically and orderly, a robust database containing all the characteristics of Free-Living Amoebas (FLAs) for each of the species described in the classification guide by taxonomic keys "A New Key to Freshwater and Soil Gymnamoebae" (1988), by Frederick C. Page, in order to employ a refined analysis to list the essential information for the classification and distinction of these microorganisms [1]. Thus, it is possible to develop a computational visualization tool, called Page's Visualization Tool (PVT), intending to elaborate a mechanism capable of improving the morphological studies of current FLAs, which are based on research and manual reading of the classification guide and therefore are not optimized.

Materials and Methods

The development of the computational tool PVT began with the construction of a database, systematically organizing the morphological descriptions for each of the species addressed in the classification keys from Page (1988) and gathering them into tables. The tabulation of the data was done in a spreadsheet with all the species described in the guide in the rows of the table, separated by class, and classified within each category explained by the author along the keys, arranged in the columns. Specific parameters were inserted for classes, orders, families, genera and, finally, species, in

order to cover as much information as possible for further refined analysis.

Once the database was ready, the construction of the software itself became possible. Considering that the database has multivariate data, since each species is related to each of the existing morphologies/characteristics and, therefore, it is a matrix with high-dimensional geometry, it was decided to use the multidimensional visualization technique of parallel coordinates to perform data analysis effectively [2-3].

Thus, from the database's .csv file, the "d3" library was used in the JavaScript language, associated with web development resources, to generate a fast and interactive program in parallel coordinates, in which the amoeba species are located in the rows and each of the morphological features in the columns.

For the development of the PVT tool, the work had the collaboration of Dr. Natália Karla Bellini, from the Butantan Institute, and Prof. Douglas Cedrim Oliveira, from the Goiano Federal Institute (IF Goiano).

Results

The tabulation of the information contained in Page's classification guide generated a database formed by 225 different species of Free-Living Amoebas, associated with 484 distinct categories of morphologies and general characteristics for such FLAs. They were distributed in information on the life stages, the fruiting body of the species that have it, the growth temperatures, the culture media used to obtain the explicit information, the description of

the lifestyle/pathogenicity and the geographic distribution of the species.

A careful investigation of the database allowed us to verify that the information about the three life stages of FLAs are fundamental for the effective morphological distinction between the species, with the trophozoite stage being the most general and complete, especially in the absence of characteristics of the cyst and flagellate stages or the presence of few of them. Data related to cysts are essential for the differentiation of species when there is a full description of them, while those associated with flagellate forms become crucial when data from cysts are not entirely sufficient. On the other hand, complementary information describing culture media, geographic distribution, lifestyle and growth temperature are not highly significant, adding unique specifications for a few species. The characteristics of the species that produce fruiting bodies, in turn, have their due relevance for differentiating the order Acrasida from the others.

In possession of this information and the database, the application of computational techniques for a preliminary analysis allowed the construction of an interactive interface for the PVT software, in order to allow the user to manually and rapidly insert the characteristics observed by microscopy in their samples and to obtain, in the best of cases, the studied species or at least the genus to which it belongs.

Conclusions

The systematic organization of the existing information in the Page guide enabled the construction of a robust and wide database containing all the descriptions brought by the author in his classification keys. The refined analysis allowed us to verify effective and relevant categories in the distinction among species, especially those associated with life stages and also the descriptive characteristics of fruiting bodies for the species that produce them. The others had an impact only for specific classifications. The accurate study of the database served as a foundation for the innovative development of the PVT software with an interactive and efficient interface. Although it has already shown effectiveness in its purpose, updates of the information are necessary, especially for the addition of both

morphological and molecular recent data to the database. Thus, it is possible to develop an innovative, optimized, effective and qualified technique to improve the studies of Free-Living Amoebas in general, relating to the morphological and molecular approach to the analyses. In this way, we believe that knowledge about FLAs can be expanded, filling the scarcity of research on the subject and reducing the socio-environmental impacts caused by this problem. In addition, such a contribution can help not only the work on this topic, but also serve as a model for the systematization of characteristics for other organisms, whether prokaryotes or eukaryotes.

References

- 1 PAGE, F.C. **A new key to freshwater and soil Gymnamoebae**: with instructions for culture. Ambleside: Freshwater Biological Association, 1988. (Culture collection of algae and protozoa). ISBN 9781871105025.
- 2 HEINRICH, J.; WEISKOPF, D. State of the art of parallel coordinates. **Eurographics** - state of the art reports. 2013. p. 95-116,. Disponível em: https://journals.de/files/heinrich_state_2013.pdf. Access in: September 5, 2022.
- 3 MUNZNER, T.. **Visualization analysis and design**. Boca Raton: CRC Press, 2014.