



STRUCTURAL TOPIC MODELING REVEALS TEMPORAL PATTERNS IN STUDIES OF METAGENOMICS APPLIED TO CONTINENTAL AQUATIC ENVIRONMENTS

Otanyel Gerlan Lima da Rocha¹, Jamira Dias Rocha¹, Rhewter Nunes¹

¹ Universidade Estadual de Goiás – UnU Iporá, Instituto Acadêmico de Ciências da Saúde e Biológicas, Iporá, Goiás, Brasil

E-mail: otanfarn@aluno.ueg.br

The identification and characterization of microbial communities in continental aquatic environments is important for the management, restoration, and sustainable use of water resources. In this context, metagenomics can be used as a tool for water quality monitoring, enabling the identification of pathogenic microorganisms, antibiotic-resistant organisms, and bioindicators of aquatic ecosystem health. Although this approach is promising, it is undergoing rapid methodological evolution due to innovations in DNA sequencing and analysis techniques, creating a gap in understanding the current state of scientific knowledge regarding its applications. Therefore, this study aims to characterize the conceptual structure of the scientific literature on the use of metagenomics in continental aquatic environments and to analyze temporal trends. The Structural Topic Modeling (STM) method was applied to the titles and abstracts of articles obtained from the Web of Science database. A total of 5,464 articles published between 2001 and 2026 were retrieved. From the article corpus, a final vocabulary of 8,029 terms was extracted and grouped into 10 topics ($K = 10$), prioritizing a balance between high likelihood and lower residuals while maintaining satisfactory semantic coherence. The analysis reveals a temporal decline in the topic focused on genome assembly from metagenomes (MAGs) in freshwater, possibly explained by the flexibility of the term “metagenomics,” which is sometimes also used in studies based on amplicon sequencing. In contrast, topics related to urban wastewater assessment and monitoring of antibiotic resistance genes (ARGs) showed expansion, indicating that these are currently among the most popular applications in the literature. Thus, this study characterizes the conceptual structure of scientific knowledge on the use of metagenomics in continental aquatic environments and reveals temporal patterns. Studies investigating patterns between amplicon-based and genome-based approaches may be promising for better characterizing the field using STM.



Keywords: *ARGs, sanitary control, wastewater*