

Generation and Curation of a Reference Sequence Database for Groundwater Organisms

Ashlyn Wandstrat, Department of Biological Sciences

Mentor: Dr. Matthew L. Niemiller, Department of Biological Sciences

Overview

One of the most biodiverse areas in the planet can be found below the Earth's surface in groundwater aquifers. These ecosystems, however, are currently being threatened increased use in agriculture and human consumption (Russo and Lall 2017). Because of these threats, many scientists have called for these ecosystems to be studied further to ensure their protection and conservation (Danielopol et al. 2003; Niemiller et al. 2018; Mammola et al. 2019). This proves challenging, however, due to the difficulty associated with reaching these habitats (Ficetola et al. 2019). Despite the difficulties that come with exploration in these areas, it is incredibly important to know what organisms live in these ecosystems in order to better protect them. In order to look into the lack of information present about groundwater organisms in commonly cited databases, I ran a R-studio program that gave the DNA sequences that are present for 1441 subterranean obligate organisms in both the BOLD and GenBank databases.

Methods

- I performed DNA extractions on numerous cave organisms that had been collected by members of Dr. Niemiller's lab from several states.

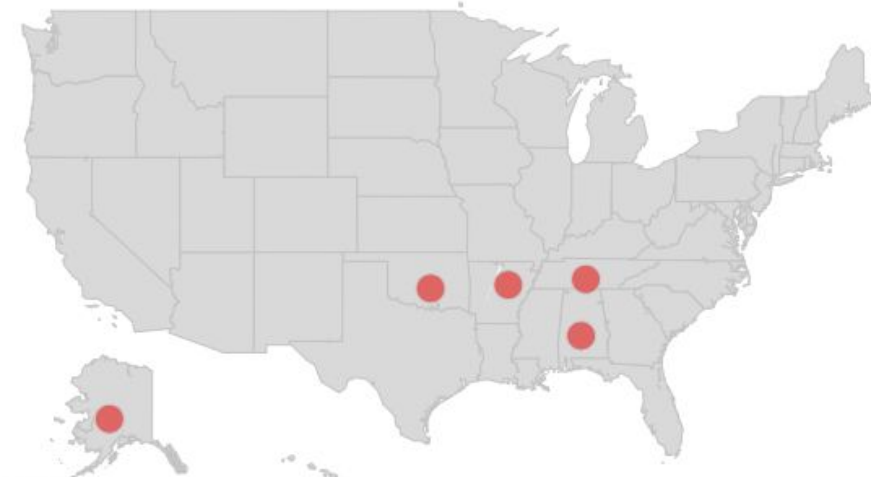


Figure 1: Map showing the states the samples in the lab were collected from: Alabama, Alaska, Arkansas, Oklahoma, and Tennessee.

- I then performed PCR on the samples to obtain the amplified DNA.
- After amplifying the DNA, I ran a gel electrophoresis to ensure that the PCR process worked correctly.
- Next, I did PCR cleanup and sent the cleaned PCR products off to be sequenced.
- After the samples were sequenced, they were analyzed and recorded.
- Then, I ran an R script to search for DNA sequences present in the BOLD and GenBank databases for subterranean organisms.

Discussion

By highlighting this lack of information on groundwater organism DNA sequences, we hope to encourage others to study and sequence more of these organisms in order to build a complete database of DNA. This would aid in ensuring that many of these species do not become extinct with the continued human interference in groundwater habitats. In total, Dr. Niemiller's lab has generated around 500 CO1 and 16S sequences for over 150 taxa across 5 states. If this research was conducted by more researchers, a more complete DNA database could be created.

References:

- Danielopol, D. L., Griebler, C., Gunatilaka, A., & Notenboom, J. (2003). Present state and future prospects for groundwater ecosystems. *Environmental Conservation*, 30(2), 104–130. doi:10.1017/S0376892903000109
- Ficetola, G. F., Manenti, R., & Taberlet, P. (2019). Environmental DNA and metabarcoding for the study of amphibians and reptiles: species distribution, the microbiome, and much more. *Amphibia-Reptilia*, 40(2), 129-148. <https://doi.org/10.1163/15685381-20191194>
- Mammola, Stefano & Cardoso, Pedro & Culver, David & Deharveng, Louis & Ferreira, Rodrigo & Fišer, Cene & Galassi, Diana & Griebler, Christian & Halse, Stuart & Humphreys, William & Isaia, Marco & Malard, Florian & Martínez, Alejandro & Moldovan, Oana Teodora & Niemiller, Matthew & Pavlek, Martina & Sofia, Ana & Reboleira, Ana Sofia & Souza-Silva, Marconi & Zgmajster, Maja. (2019). Scientists' Warning on the Conservation of Subterranean Ecosystems. *BioScience*. 69. 641–650. 10.1093/biosci/biz064.
- Niemiller, Matthew & Taylor, Steven & Bichuette, Maria Elina. (2018). Conservation of Cave Fauna, with an Emphasis on Europe and the Americas: Analysis and Synthesis. 10.1007/978-3-319-98852-8_22.
- Russo, T., Lall, U. Depletion and response of deep groundwater to climate-induced pumping variability. *Nature Geosci* 10, 105–108 (2017). <https://doi.org/10.1038/ngeo2883>

Acknowledgements

I would like to thank the UAH Honors College for funding this project. I would also like to thank Dr. Matthew L. Niemiller and Jared Higgs for their contributions and assistance on this project.

Results

- In the BOLD database, nearly 90% of the organisms that were included did not have DNA sequences present.
- The GenBank database had slightly more organisms with DNA sequences present, with nearly 82% of the included organisms not having sequences present.

Groundwater Organism DNA Sequences in BOLD Database

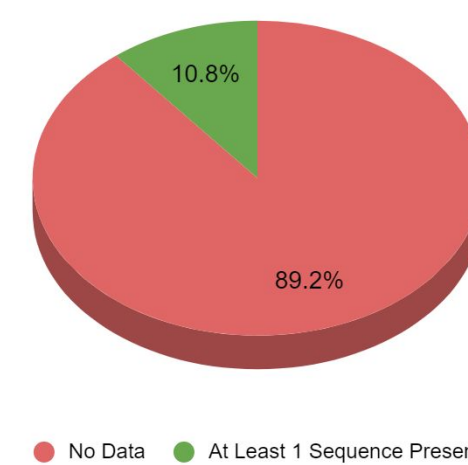


Figure 2: Pie chart showing the percent of organisms that have at least 1 DNA sequence recorded in the BOLD database.

Groundwater Organism DNA Sequences in GenBank Database

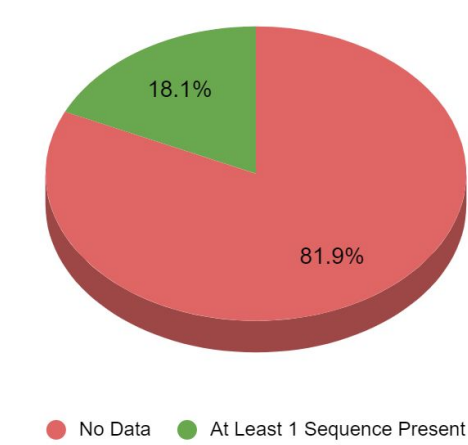


Figure 3: Pie chart showing the percent of organisms that have at least 1 DNA sequence recorded in the GenBank database.