

# Survival Strategies of High GC-Content Microorganisms in Oligotrophic Deep Groundwater

Maryam Rezaei Somee<sup>‡</sup>, Carolina González-Rosales<sup>‡</sup>, Stephanie Turner<sup>§</sup>, Stefan Bertilsson<sup>||</sup>, Mark Dopson<sup>‡</sup>, Maliheh Mehrshad<sup>||</sup>

<sup>‡</sup> Centre for Ecology and Evolution in Microbial Model Systems (EEMiS), Linnaeus University, Kalmar, Sweden

<sup>§</sup> Department of Forest Mycology and Plant Pathology, Swedish University of Agricultural Sciences, Uppsala, Sweden, Uppsala, Sweden

<sup>||</sup> Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences (SLU), Box 7050, 75007, Uppsala, Sweden, Uppsala, Sweden

Corresponding author: Maliheh Mehrshad ([maliheh.mehrshad@slu.se](mailto:maliheh.mehrshad@slu.se))

## Abstract

Deep groundwaters are among the most energy and nutrient-limited ecosystems on the planet. The limited resources are mainly due to the absence of photosynthesis-driven primary production (Kadnikov et al. 2020). These ecosystems do however host phylogenetically diverse and metabolically active microorganisms from all domains of life plus viruses (Holmfeldt et al. 2021, Mehrshad et al. 2021). In this study, we used a large metagenomic dataset generated over the last eight years from the Äspö Hard Rock Laboratory (Äspö HRL) in Sweden and drill holes in Olkiluoto Island, Finland. This dataset, termed the “Fennoscandian Shield Genomic Dataset” (FSGD), contains metagenome-assembled genomes (MAGs) and single-cell amplified genomes (SAGs). Previous studies on this dataset have shown that reciprocal symbiotic partnerships and efficiency of energy metabolism define the core microbiome of these deep groundwaters (Mehrshad et al. 2021). Studies on different marine and freshwater ecosystems show that oligotrophic environments host streamlined genomes with lower GC content. However, it is not known how the low carbon and energy availability in deep groundwaters affect the microbial community regarding their genome size and GC content. To address this, we used the FSDG to study the distribution of genome size and GC content among bacterial and archaeal genomes in Fennoscandian Shield deep groundwaters. We further disentangled the prevalent metabolic strategies in these genomes that is being used to support their carbon and nitrogen demands for replication and survival.

A total of 1990 MAGs/SAGs with a completeness of  $\geq 50\%$  and  $< 5\%$  contamination were recovered from 43 metagenomic datasets. The taxonomy of the MAGs/SAGs was assigned using the GTDB-tk (Chaumeil et al. 2022) and the GC content and genome size of MAGs/SAGs were calculated. MAGs/SAGs were also functionally annotated to investigate the genome-encoded functional potential. To survey the preference for different metabolic pathways and metabolic cross-feeding, the C-fixation pathways were subdivided

into 82 routes that lead to the production of intermediate compounds (e.g., formate, pyruvate, oxaloacetate, etc.). A similar analysis for nitrogen acquisition pathways and CO<sub>2</sub>-capturing enzymes was performed and the presence/absence of these metabolic modules was investigated in representative MAGs/SAGs from various deep groundwater types to explore the microbial community's metabolic interaction.

The GC content and estimated genome size (EGS) of the recovered MAGs/SAGs were linearly correlated, suggesting that higher genome-level GC content is associated with larger genome size. The most common taxa among high GC (≥50%) content MAGs/SAGs were affiliated with the phyla Proteobacteria, Desulfobacterota, Actinobacteriota, Chloroflexota, and Patescibacteria. Analysis of metabolic modules in the environmental context revealed that high GC content MAGs constituted the main primary producers in all investigated groundwaters. Among the seven investigated C-fixation pathways, routes leading to the production of intermediates of reductive tricarboxylic acid (rTCA) were significantly enriched in high GC-content genomes. Both carboxylation/reduction first CO<sub>2</sub>-capturing enzymes were significantly enriched in MAGs with high GC contents. Of the various carbon fixation strategies, rGly and rTCA are the most energy-efficient and consume one and three ATP molecules, respectively (Bar-Even et al. 2011), matching the high prevalence of rTCA in oligotrophic deep groundwater. High GC content genomes featured metabolic pathways to generate key precursors for synthesizing purine/pyrimidine and amino acids. As GC base-pairs require one more nitrogen than AT base-pairs, prokaryotes with a higher GC content need proportionally more nitrogen for genome replication. In agreement with this, high GC content MAGs/SAGs from the FSGD feature additional nitrogen uptake systems including ammonium permease and N-fixation genes.

High GC content MAGs/SAGs appear to be metabolically versatile and capable of acquiring nutrients via different carbon and nitrogen fixation pathways along with various C-scavenging strategies for substrate uptake to acquire energy and survive in oligotrophic conditions. Although their larger genome size and higher GC content entail higher expenses for replication and maintenance, their presence within the community of deep groundwater is supported by the "Black Queen Theory". According to this theory, a cell containing an expensive metabolic pathway is maintained within the community as it provides a service that other members cannot.

## **Keywords**

deep groundwater, GC-rich genomes, carbon fixation, nitrogen fixation

## **Presenting author**

Maryam Rezaei Somee

## Presented at

ISEB-ISSM 2023

## Hosting institution

Linnaeus University

## Conflicts of interest

The authors have declared that no competing interests exist.

## References

- Bar-Even A, Noor E, Milo R, et al. (2011) A survey of carbon fixation pathways through a quantitative lens. *Journal of Experimental Botany* 63 (6): 2325-2342. <https://doi.org/10.1093/jxb/err417>
- Chaumeil P, Mussig AJ, Hugenholtz P, Parks DH (2022) GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* 38 (23): 5315-5316. <https://doi.org/10.1093/bioinformatics/btac672>
- Holmfeldt K, Nilsson E, Simone D, Lopez-Fernandez M, Wu X, de Bruijn I, Lundin D, Andersson A, Bertilsson S, Dopson M, et al. (2021) The Fennoscandian Shield deep terrestrial virosphere suggests slow motion 'boom and burst' cycles. *Communications Biology* 4 (1). <https://doi.org/10.1038/s42003-021-01810-1>
- Kadnikov V, Mardanov A, Beletsky A, Karnachuk O, Ravin N, et al. (2020) Microbial Life in the Deep Subsurface Aquifer Illuminated by Metagenomics. *Frontiers in Microbiology* 11 <https://doi.org/10.3389/fmicb.2020.572252>
- Mehrshad M, Lopez-Fernandez M, Sundh J, Bell E, Simone D, Buck M, Bernier-Latmani R, Bertilsson S, Dopson M, et al. (2021) Energy efficiency and biological interactions define the core microbiome of deep oligotrophic groundwater. *Nature Communications* 12 (1). <https://doi.org/10.1038/s41467-021-24549-z>