

# Within and between gene variants: tracking for potential targets for populational linkage according to the metagenomic profile in a changing freshwater environment

Marcele Laux, Ricardo A. Vialle, José Miguel Ortega, Alessandra Giani

*Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais*

The *Microcystis* population structure of a tropical freshwater reservoir under bloom development was analyzed through metagenomic whole genome shotgun in order to find targets of interaction among population members in a changing environment. The metagenomic profile was followed over a temporal and spatial scale and pointed to a 5-fold increase in *Microcystis* assignments after the beginning of biomass increase and a similar 5-fold increase from the dam towards the upstream sampling station. Sixteen strains were assigned in the samples and *M. aeruginosa* NIES-843 was the most abundant and frequently distributed. A two-step recruitment was performed to analyze variable and conserved genetic contents of the NIES-843 sequence-discrete population. Thirteen hypervariable regions were identified, presenting lower alignment attributes and positional symmetry among the samples. A read mapping and population Single Nucleotide Variant (SNV) calling was then performed focusing on those regions, to track the gene variability within (above 99% similarity to NIES-843) and between (95-98% similarity) the mapping, since the variations at nucleotide level can be detected even among distinct populations. The variants were called according to GATK pipeline, and classified according to protein transcription impact. Only variants in coding regions were considered. The energetic metabolism seems to be under strong pressure considering the variation between populations, given that several of the variants identified are involved in regulation of the photosynthesis and respiration systems. The PSI assembly protein Ycf3, OrfB transposases family and Nudix hydrolases family were among such set of genes, which could be potentially related to the growing biomass of the *Microcystis* population over the sampling period. Similarly, the variants identified within the population were mostly related to regulation processes, noteworthy the PatA family of microbial two-component response regulators, and a transcriptional regulator Pirin-like protein, indicating potential functional keys for intraspecific dynamics. Additionally, sixty one housekeeping genes presented high impact variants, most of them with variant allele frequency of 1.0 and distinct occurrence according to each strain. We highlight the Thioredoxin gene *trxA*, essential for photosynthetic growth, which presented a generally low but increasingly depth towards the development of the bloom event, showing high impact variants in most reference genomes. The *Microcystis* population presented a mosaic-like genomic structure with distinct gene contribution among the strains. The “within” and “between” non-synonymous SNVs were mostly related to transposases and regulators, potentially associated to energetic demands and growth conditions, considering that those differentially aligned target genes were mapped mostly or even exclusively in samples from bloom events.