



**Figure S1. Top ten most expressed genes in each lake.** Coding regions from reference genomes and metagenome assemblies were clustered at 97% sequence similarity, and the longest coding region was chosen as the representative sequence. Metatranscriptomic reads were mapped to these representative sequences. Annotations and classifications are derived from the sequence to which each read mapped. Read counts were summed across each lake and log transformed for visualization. Panels A-C show the top ten most expressed genes in each lake, while panels D-F show the top ten annotated genes not directly related to primary production (photosynthesis or carbon fixation) in each lake. Phylum is indicated by color in all panels. Proteobacteria were split into classes due to the high diversity of this phylum.