

Fig. S1. A tiled display of an emergent self-organizing map (ESOM) based on the tetranucleotide frequency (TNF) of the 19 Verrucomicrobia MAGs. TNF was calculated with a window size of 5 kbp, with each dot on the ESOM representing a 5-kbp fragment (or a contig if its length is shorter than 5 kbp). Dots (i.e. fragments) are colored according to MAGs. A numeric ID is assigned to each MAG, and IDs from Mendota are labeled in black and IDs from Trout Bog labeled in white. A red outline was drawn to indicate the clustering of MAGs from Mendota on the ESOM.

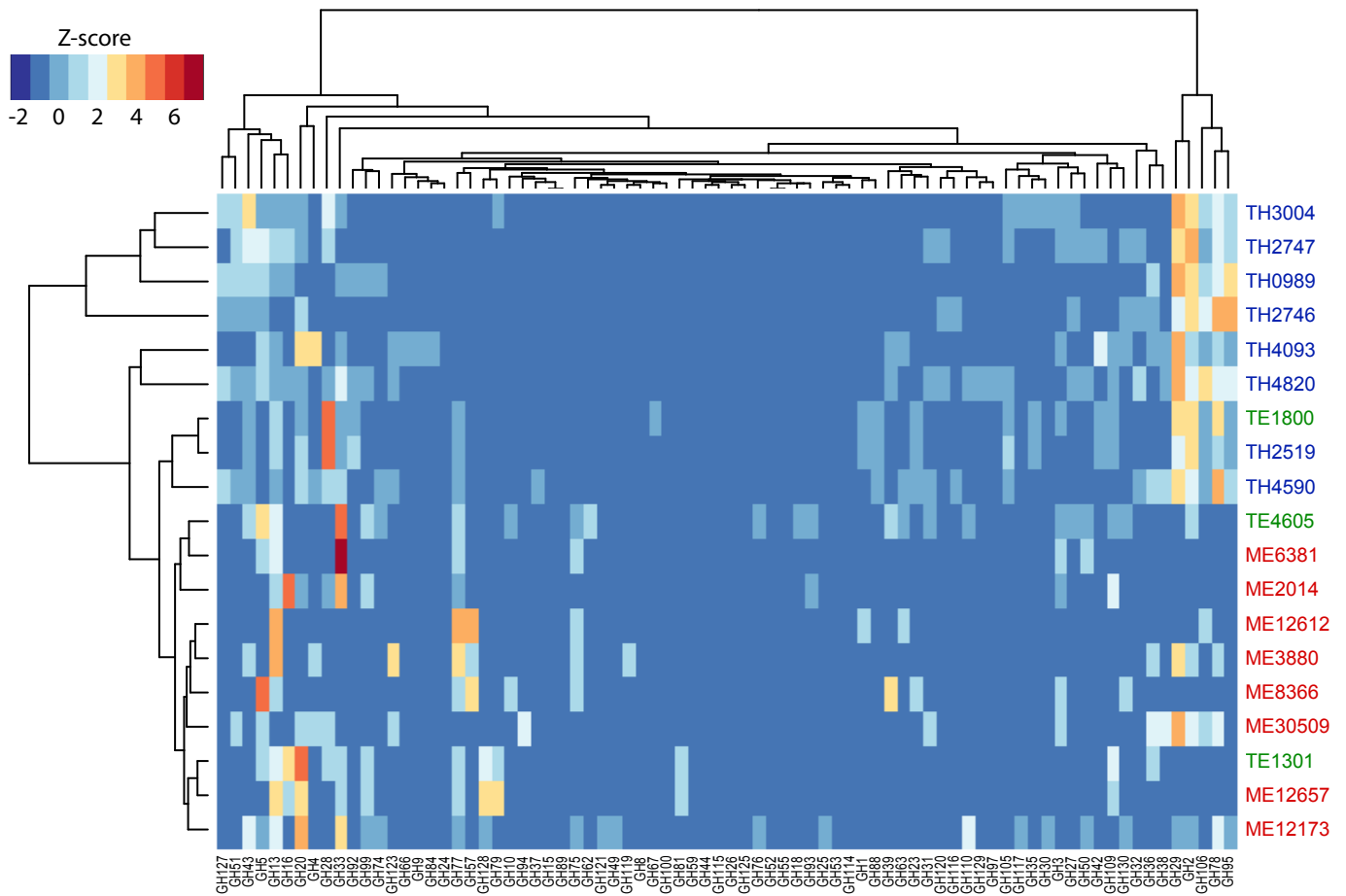
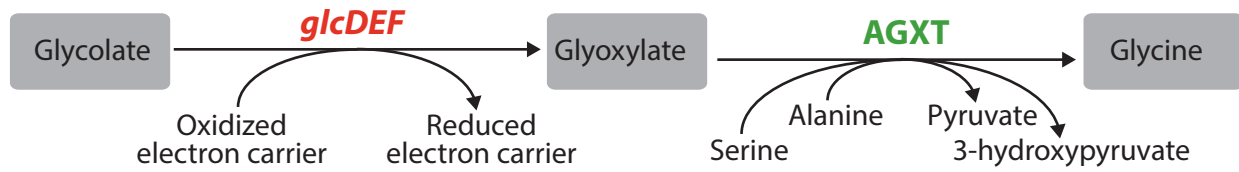
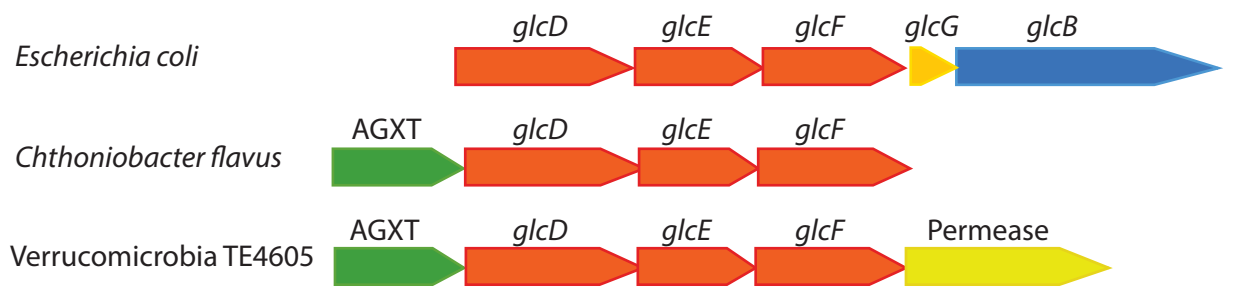


Fig. S3. Heat map based on GH abundance profile patterns showing the clustering of MAGs by different lakes.



glcDEF: subunits of glycolate oxidase

glcG: Uncharacterized *glc* operon protein GlcG

glcB: Malate synthase G

AGXT: Alanine-glyoxylate transaminase / serine-glyoxylate transaminase / serine-pyruvate transaminase

Permease: L-lactate permease family to which GlcA (glycolate permease) belongs to.

Fig. S5. Comparison of glycolate oxidase gene operons in *E. coli*, *C. flavus* and TE4605.

