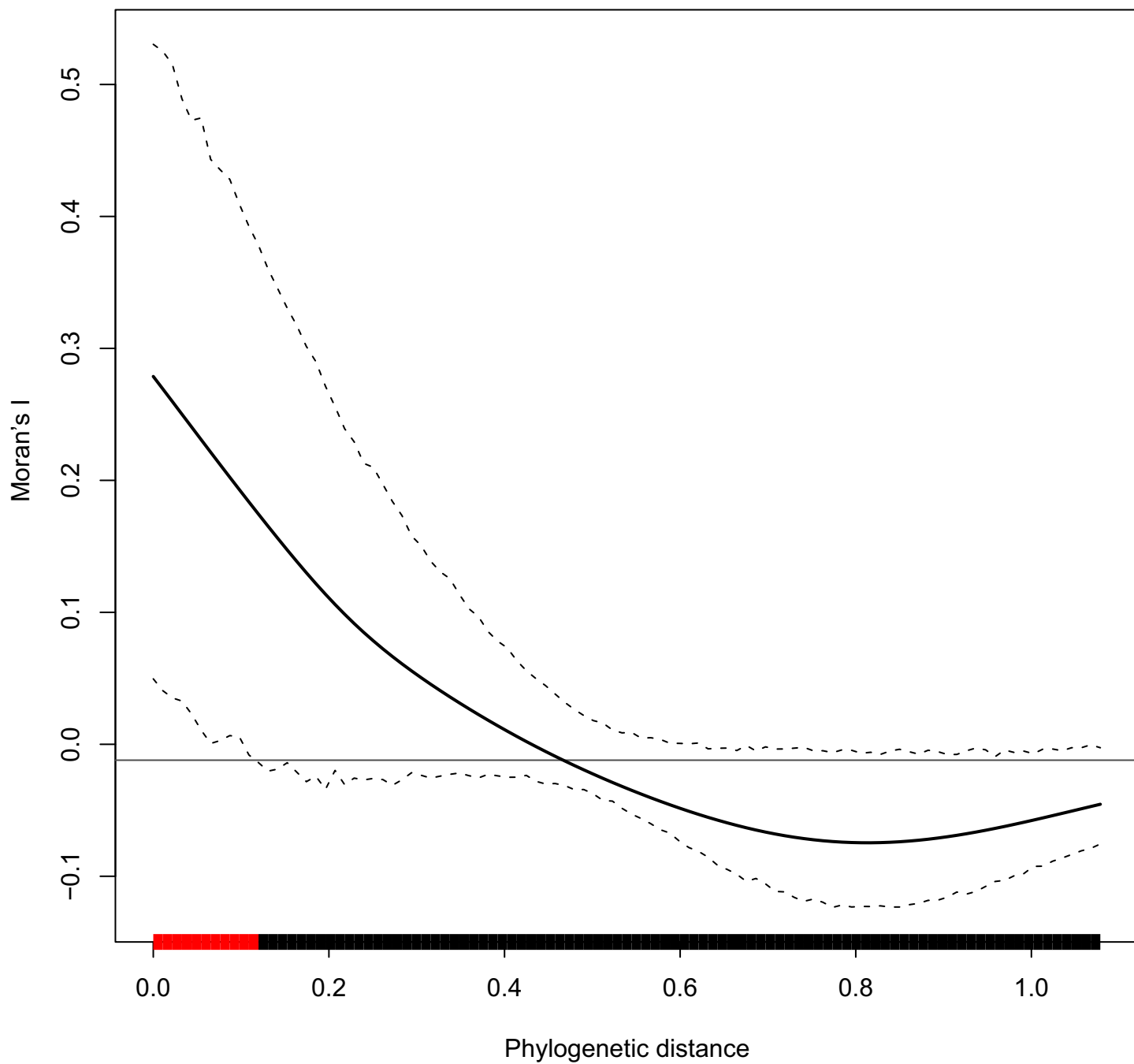


a

Phylogenetic correlogram



b

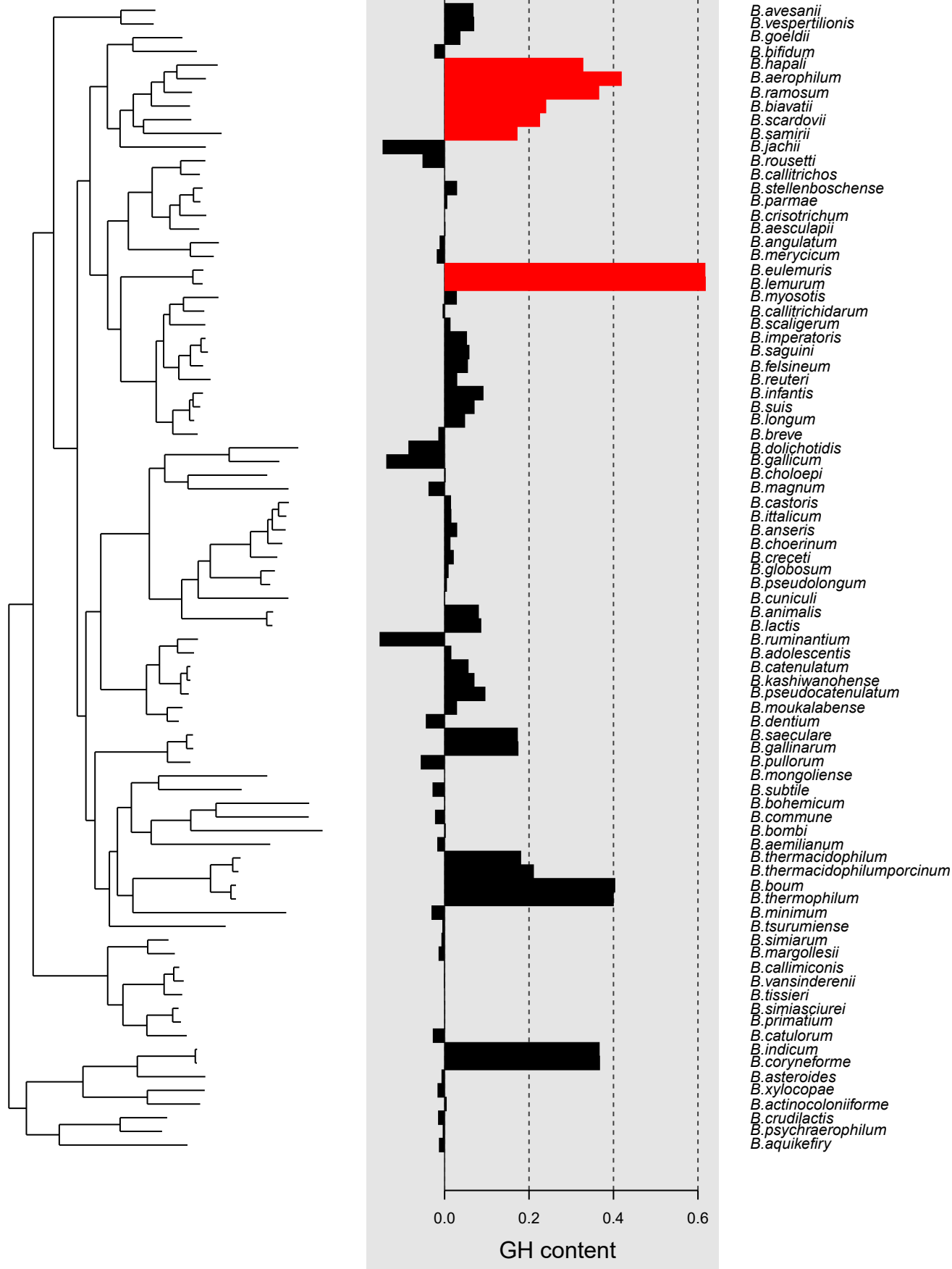


Figure S3. (a) Phylogenetic correlogram based on GH content. No significant autocorrelation was observed above the phylogenetic distance of 0.1. The dash lines represent the lower and upper confidence intervals and solid line represents the Moran's I index of autocorrelation. The colored horizontal bars at the bottom shows the significance of autocorrelation: red- significant positive autocorrelation, blue- significant negative autocorrelation and black- no autocorrelation; **(b)** Local Moran's index values for GH content for each type strain. The clades highlighted in red shows the presence of significant phylogenetic signal for p-value < 0.01.