Supporting Information

Figure S1: Nonpareil curves for the 22 metagenomes. The plot displays the fitted models of the Nonpareil curves. The horizontal dashed lines indicate 100 (gray) and 95% (red) coverage. The empty circles indicate the size and estimated average coverage of the datasets, and the lines after that point are projections of the fitted model.

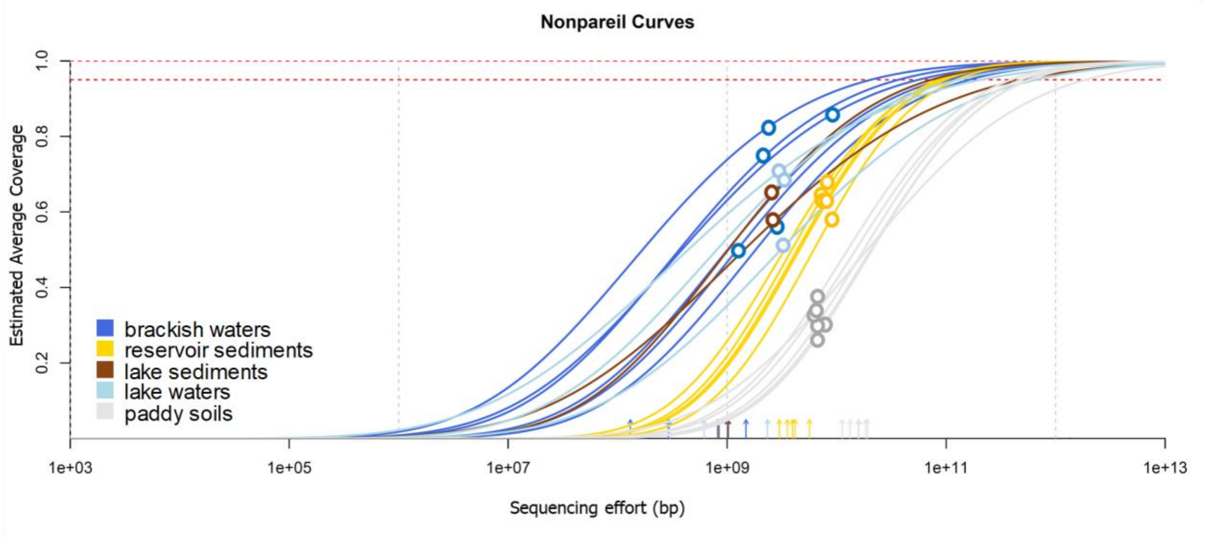

Figure S2. Distribution of *hgcA* genes in the 22 metagenomes recovered using the co-assembly ´c´ and the single assembly ´s´ methods and applying the three stringency cutoffs defined in this manuscript for the definition of *hgcA* genes. Abundance values were calculated as *hgcA* coverage values normalized by *rpoB* normalized values. Colors denote taxonomic affilitations of *hgcA* genes.

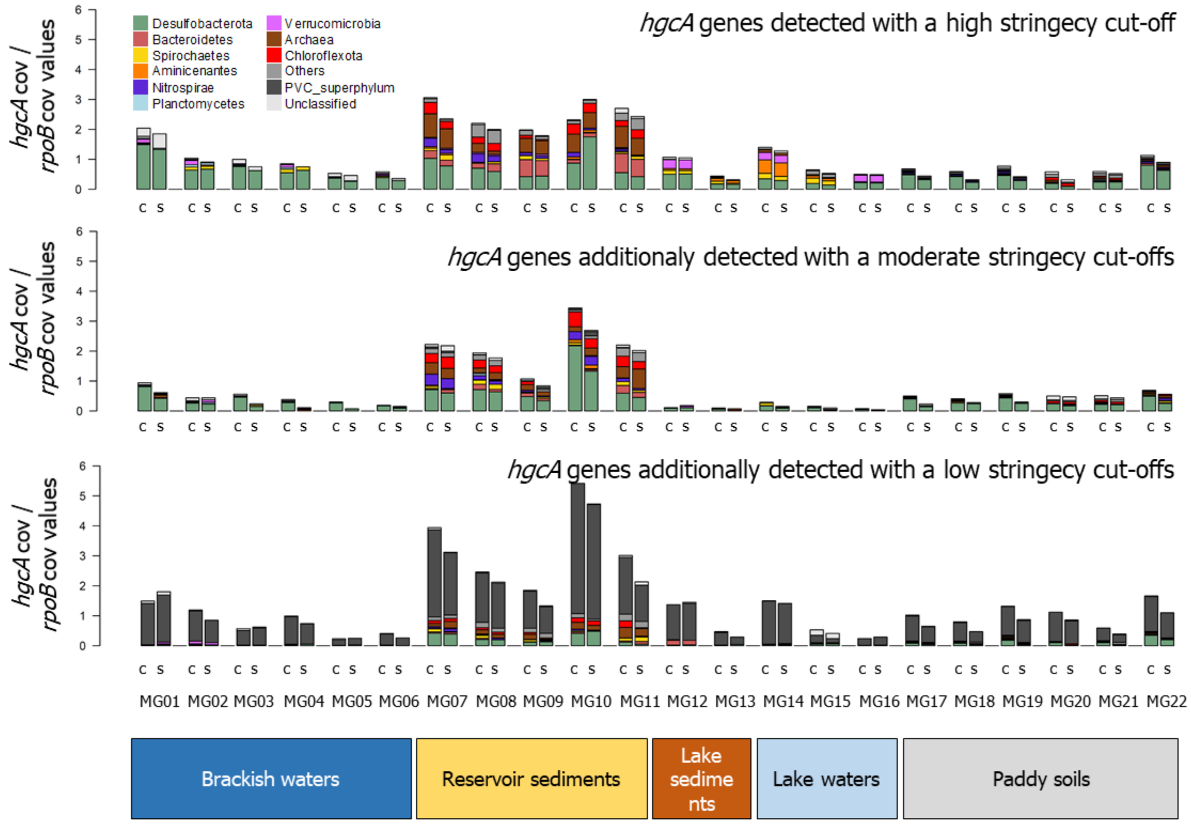

Figure S3: Violin boxplots showing, for each metagenome, the difference in *hgcA* sequence length distribution comparing the outputs of the co-assembly and the single assembly approaches.

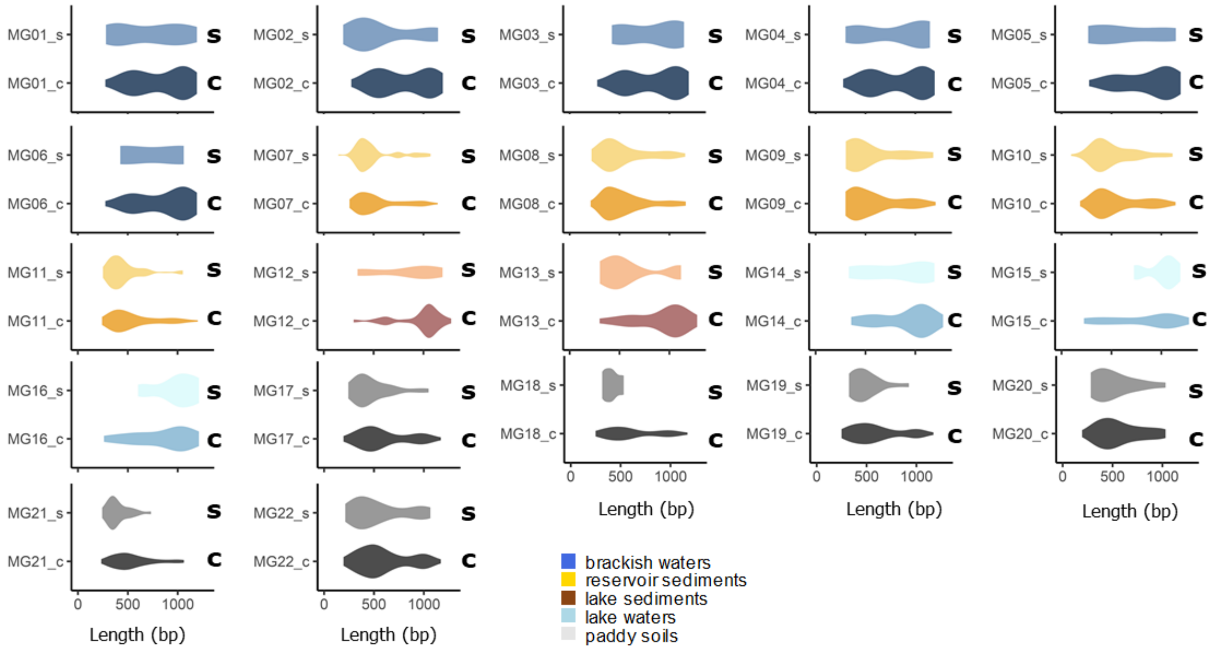

Figure S4: Distribution of *hgcA* genes in the 22 metagenomes with the co-assembly (c) and the single assembly (s) methods with different normalization methods