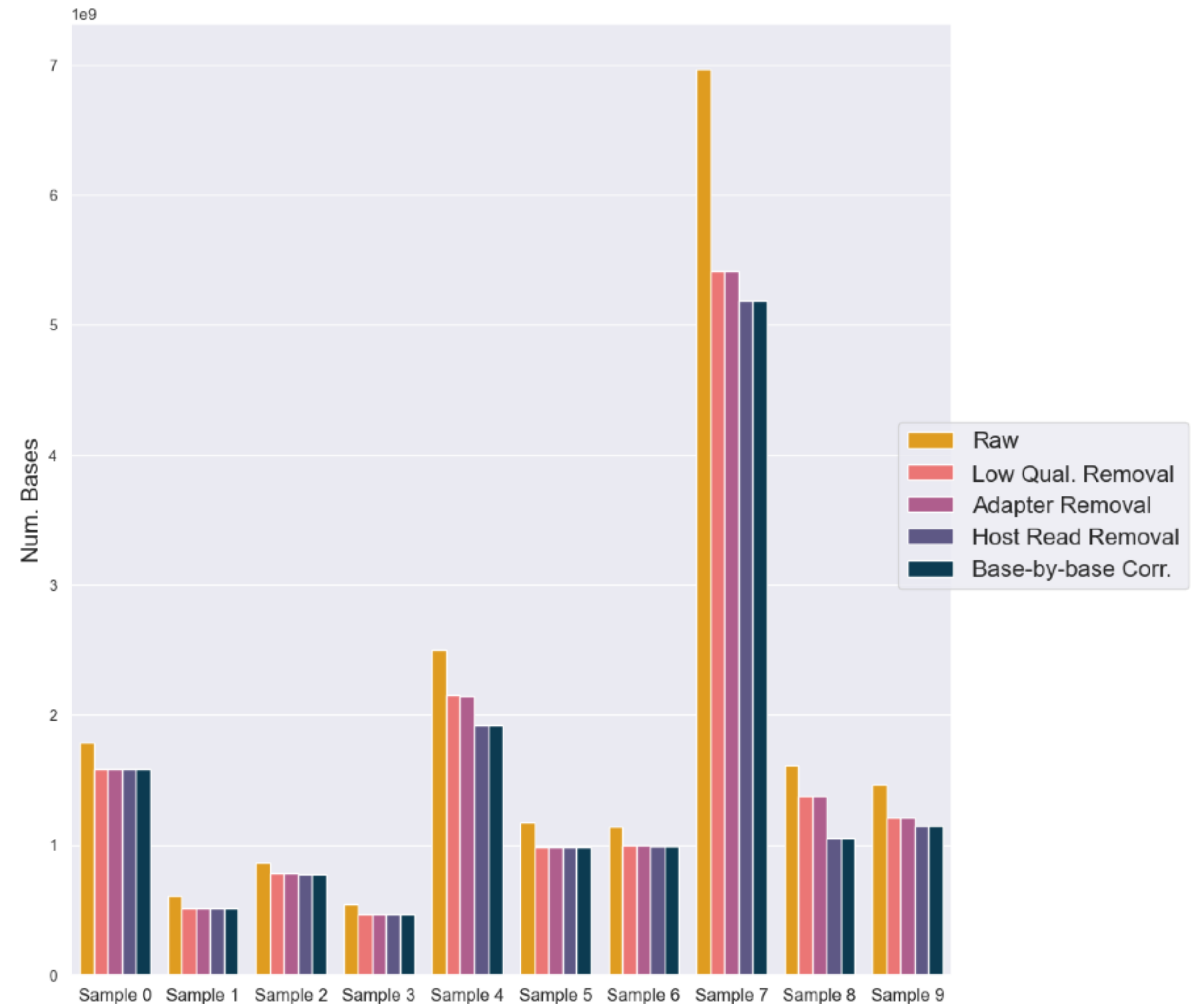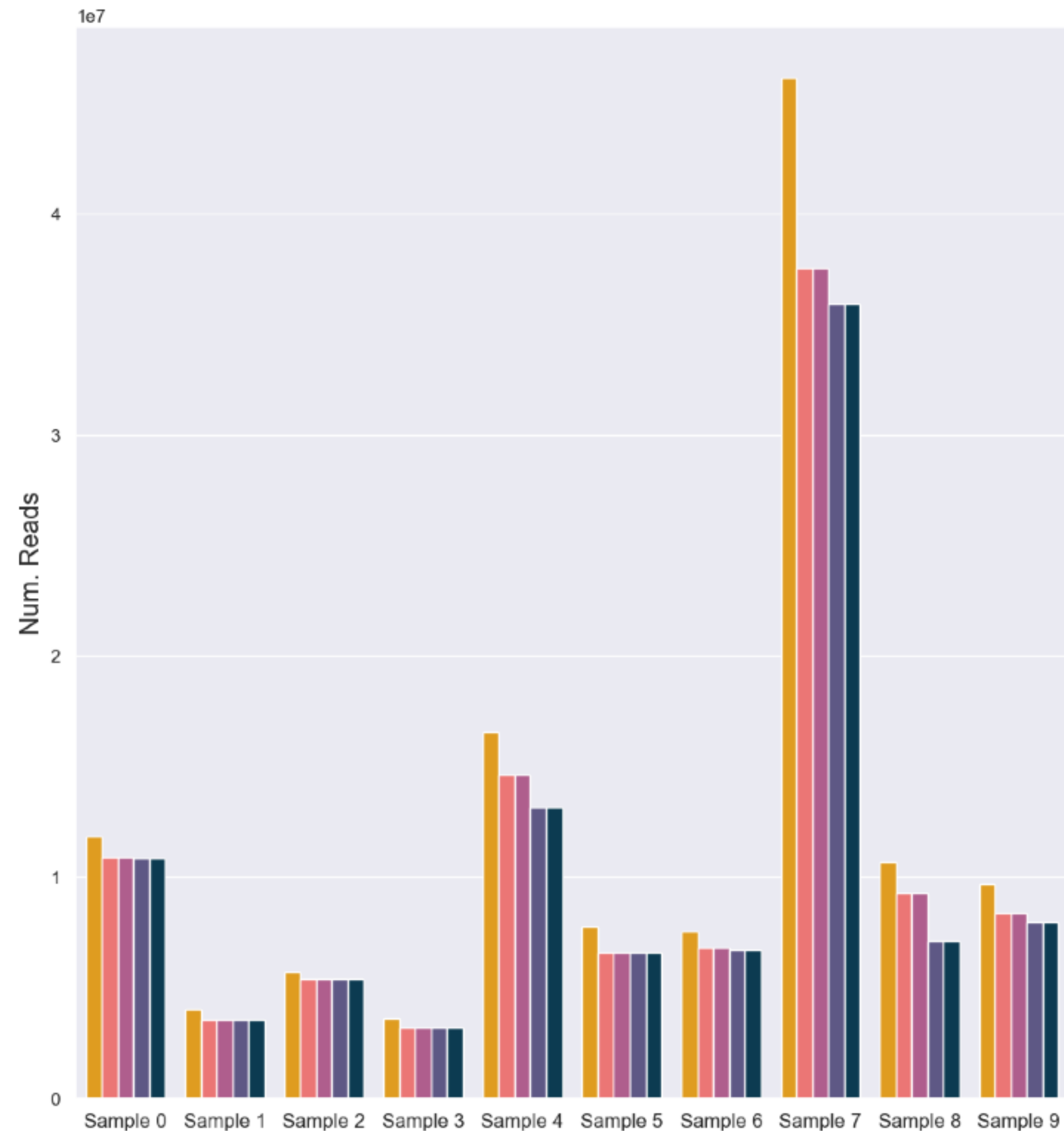# Quality control and preprocessing of sequencing data. **(A)** Number of reads and **(B)** Bases retained

Relationship between sequencing dataset size and assembly statistics across samples