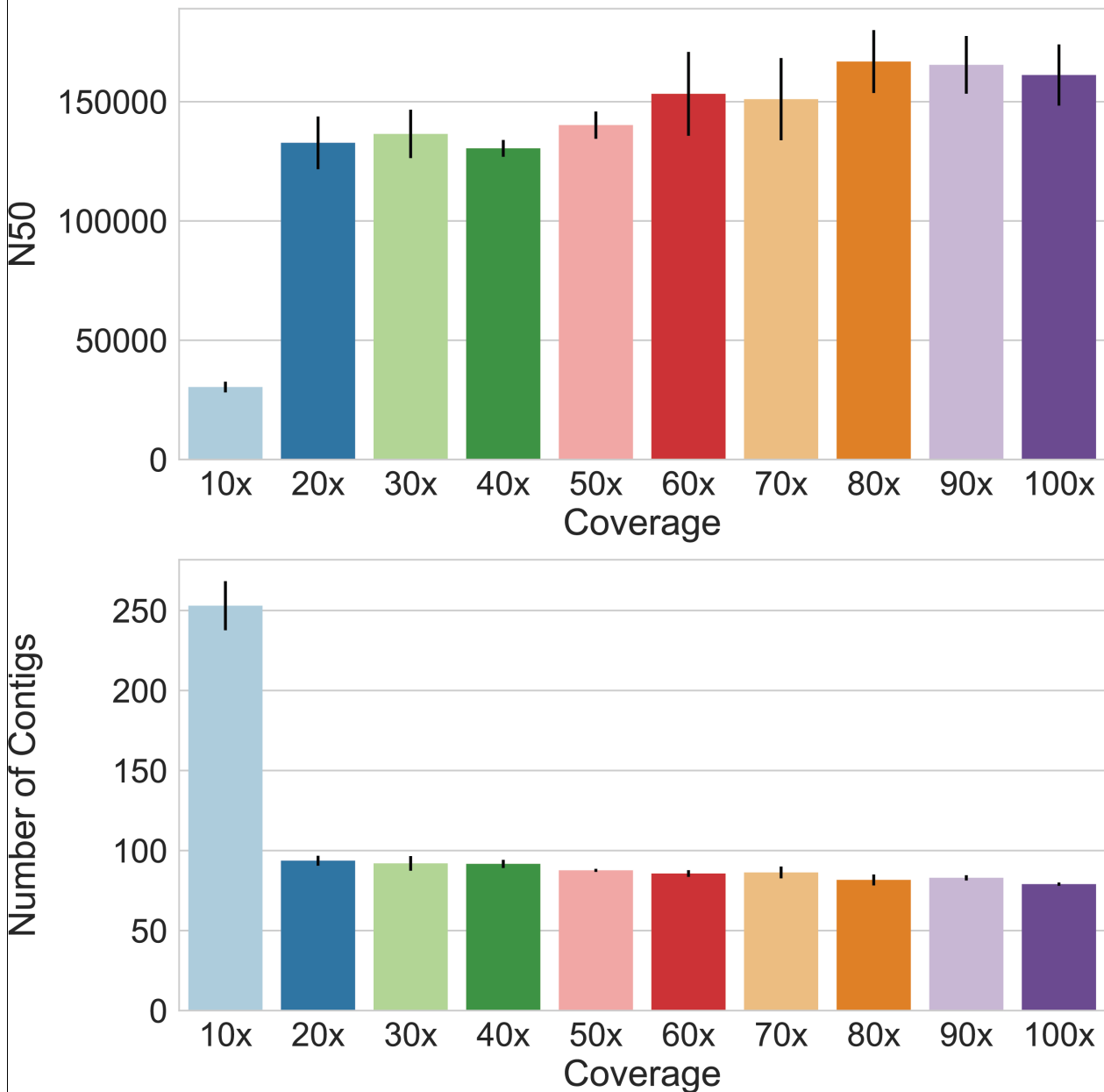


# Assembly Statistics Following Randomized Sampling to Desired Coverage



## Supplementary Figure 1

### Resulting Assembly Statistics at Various Coverage

Resulting Assembly Statistics at Various Coverage. To investigate the effect of coverage on overall assembly statistics of N50 and Number of Contigs, we randomly sampled reads of the BOP27 strain, which has been sequenced to extremely high coverage (400X), enabling this analysis. Analyzing the resulting assemblies at coverages ranging from 10X to 100X, we see from comparing the metrics that at 70X the assembly quality mostly saturates and as such we recommend included genomes have at least this much coverage.