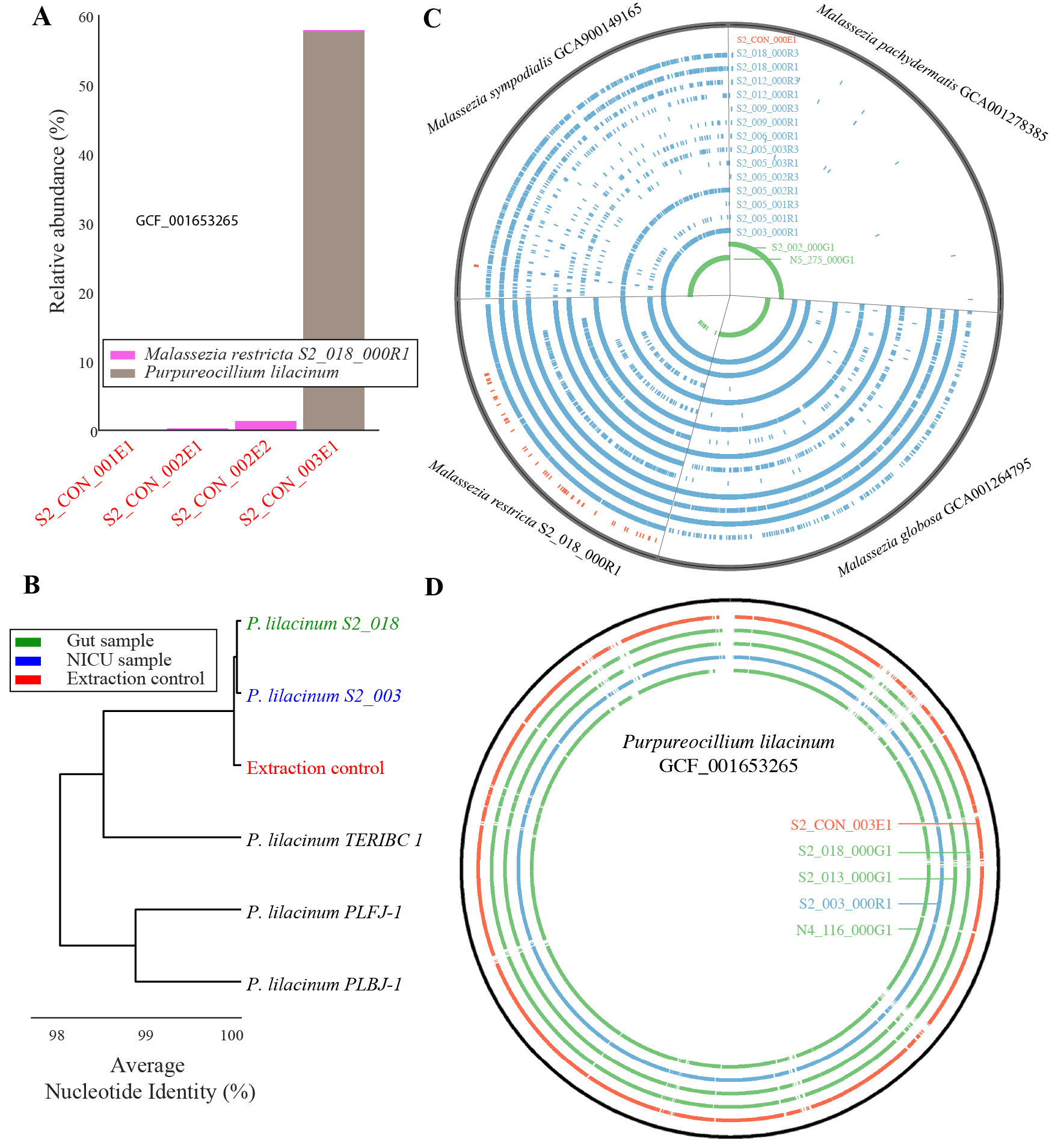
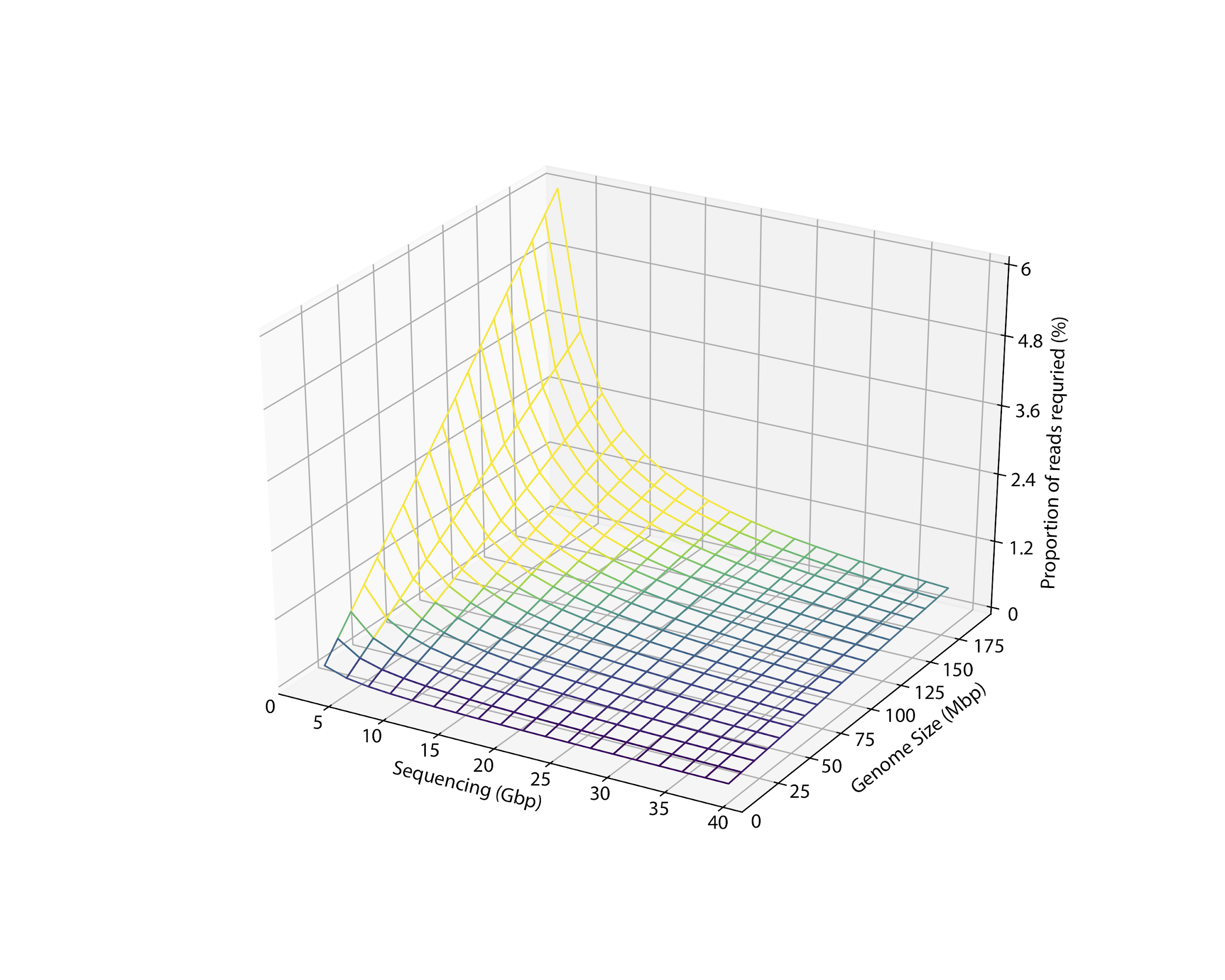
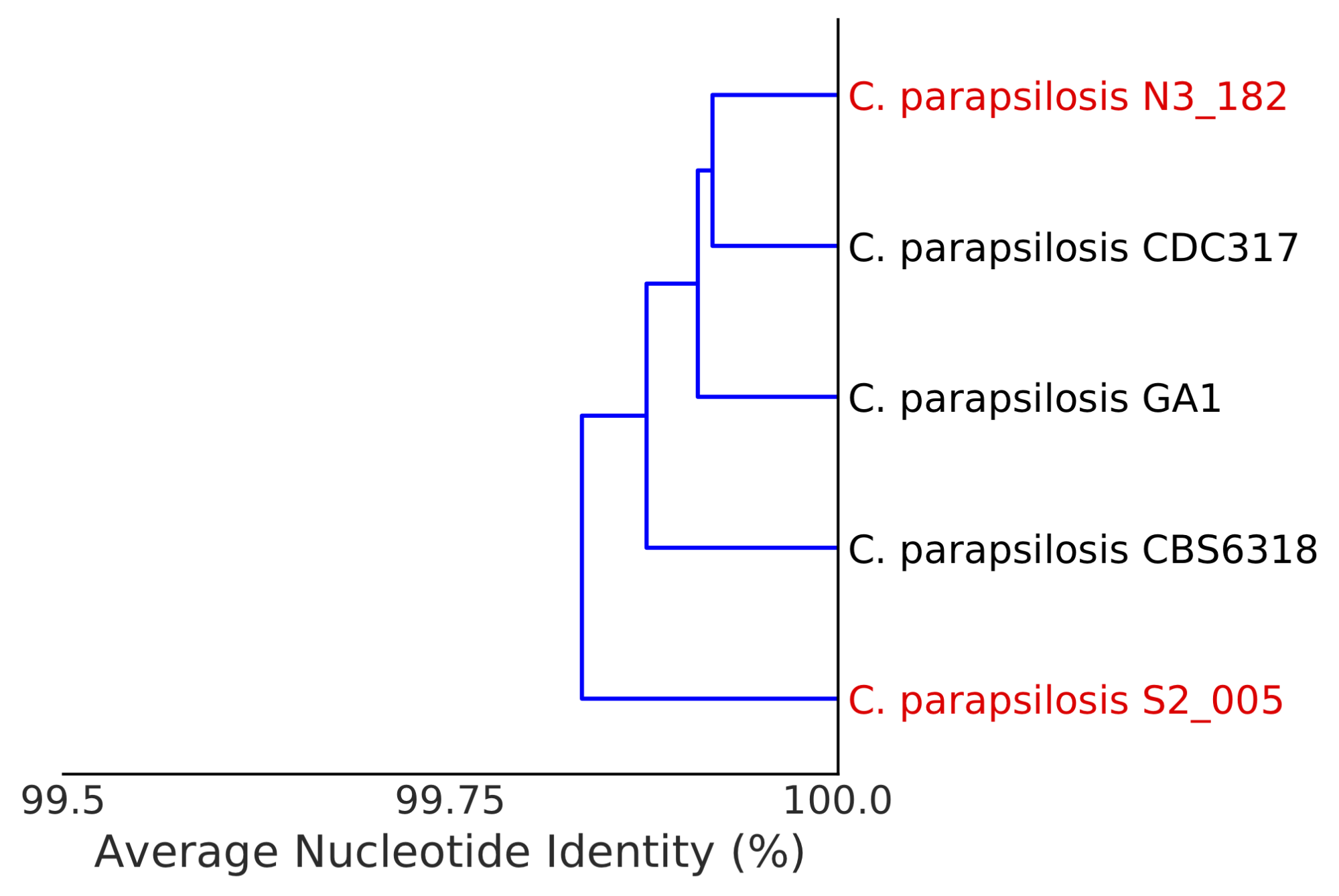
****

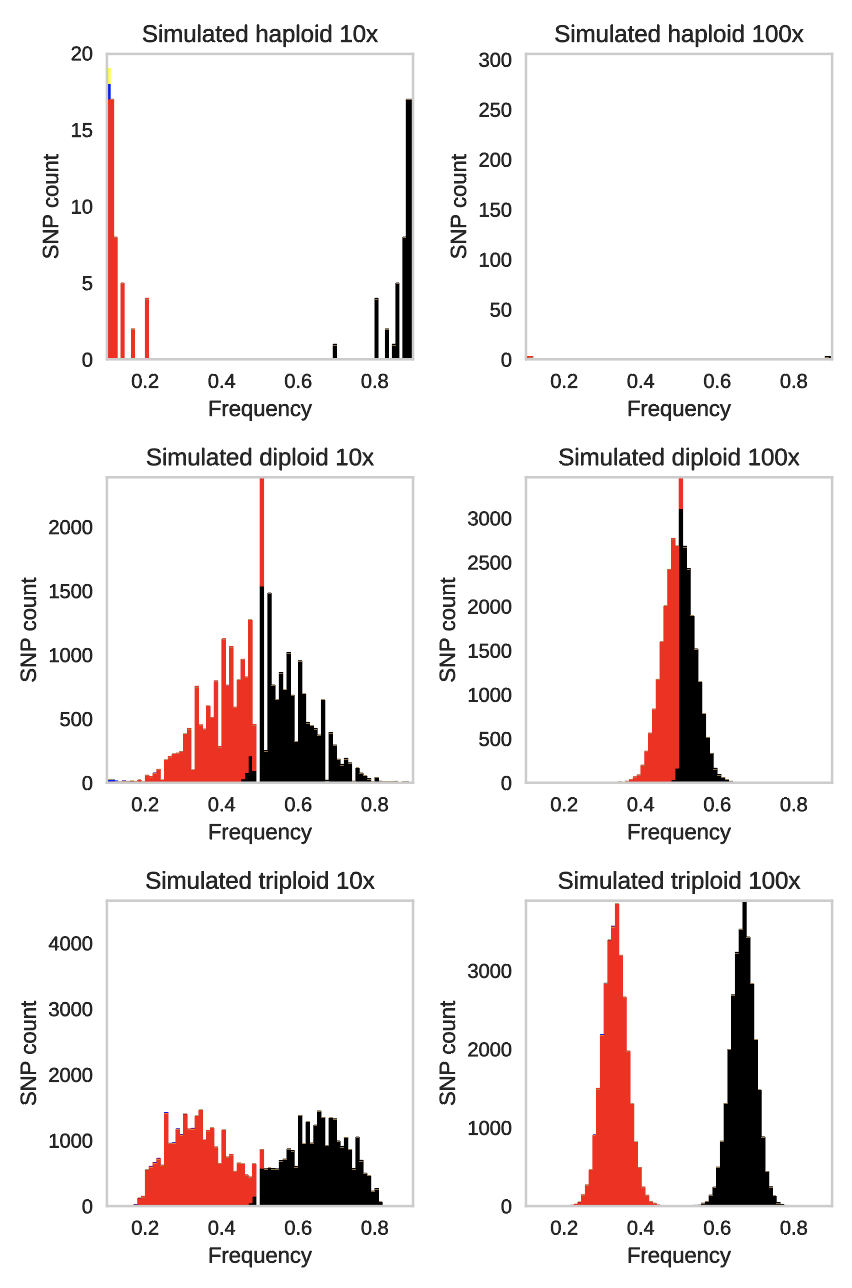
**Supplemental Figure S1.** Fungal contaminants are present in negative extraction controls. (A) Relative abundance of eukaryotes in four sequenced extraction controls (based on read mapping). (B) *P. lilacinum* sequences from the extraction control (red) closely resemble sequences recovered from gut and room samples (blue), and are distinct from publically available genomes (black). (C, D) Each ring shows the breadth of coverage across (C) four different Malassezia genomes or (D) a *Purpureocillium lilacinum* reference genomes for an individual sample. Red, blue, and green rings are extraction controls, NICU room samples, and premature infant guts samples respectively. Each colored tick represents a 10 kb window in which the breadth of coverage of that genome is at least 50%.



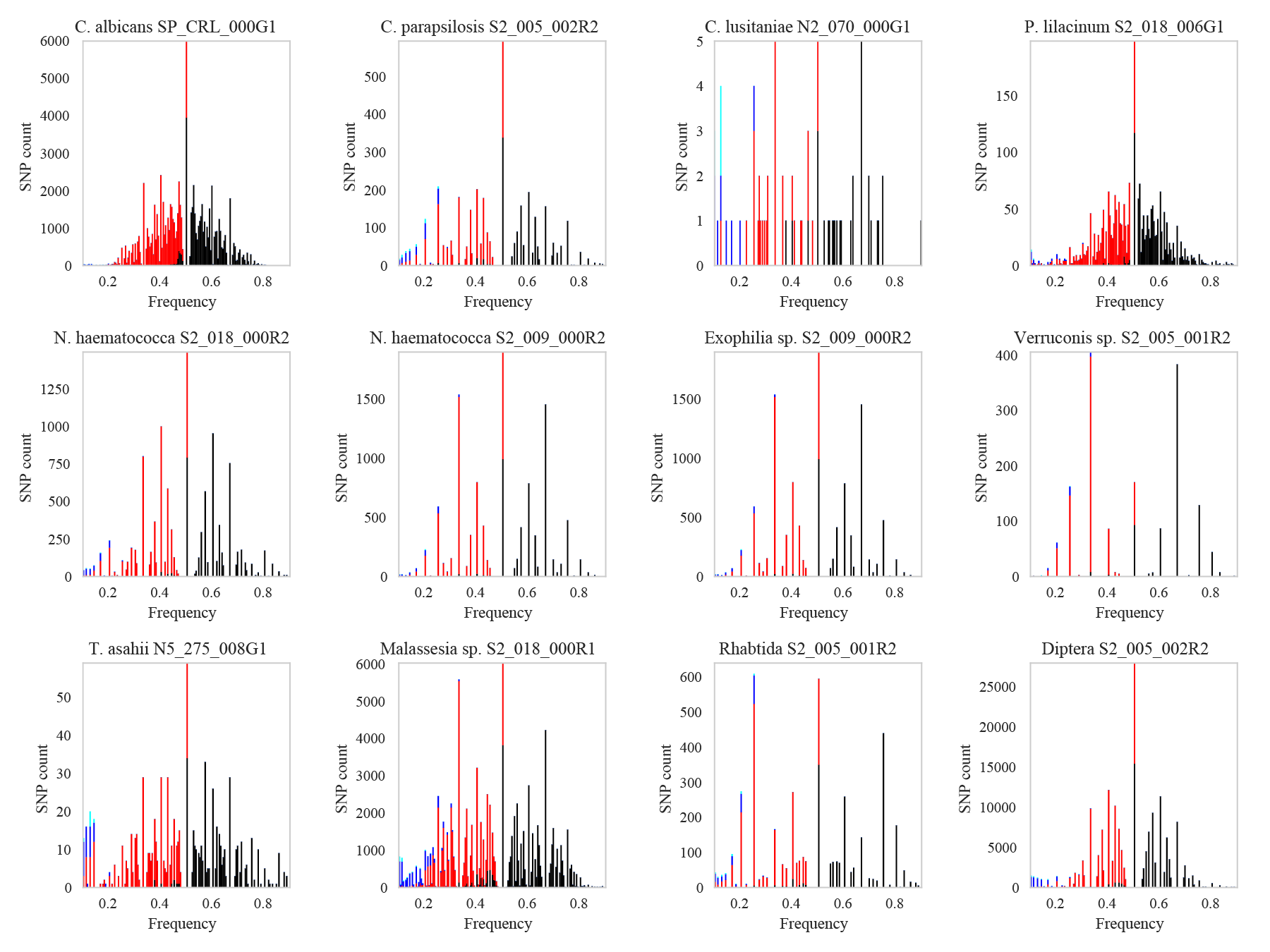
**Supplemental Figure S2.** The sequencing depth and relative abundance needed to detect eukaryotic genomes of various lengths at 1x coverage.



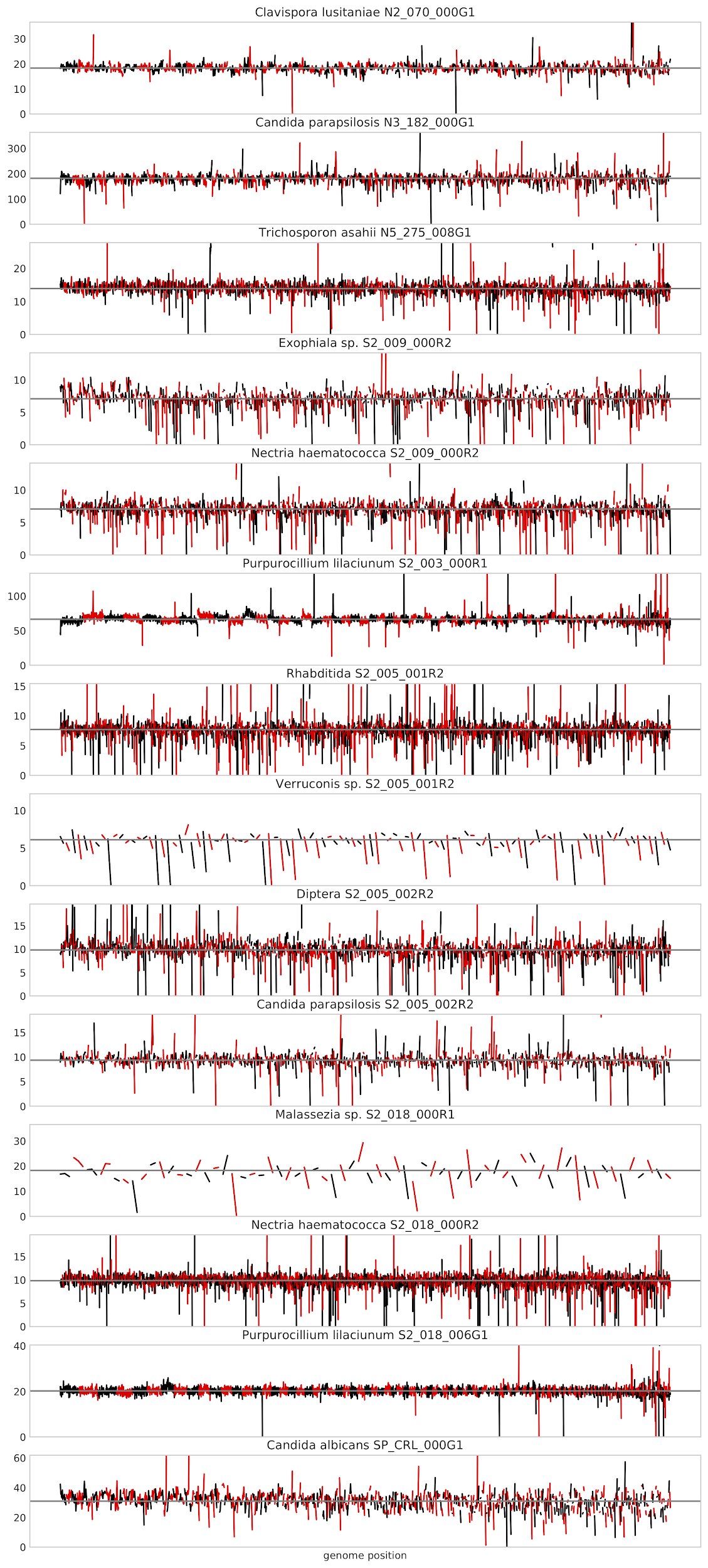
**Supplemental Figure S3.** Dendrogram of reference and *de novo* assembled *C. parapsilosis* genomes based on genome-wide average nucleotide identity (ANI). Genomes assembled from this study are in red, and public reference genomes are in black.

****

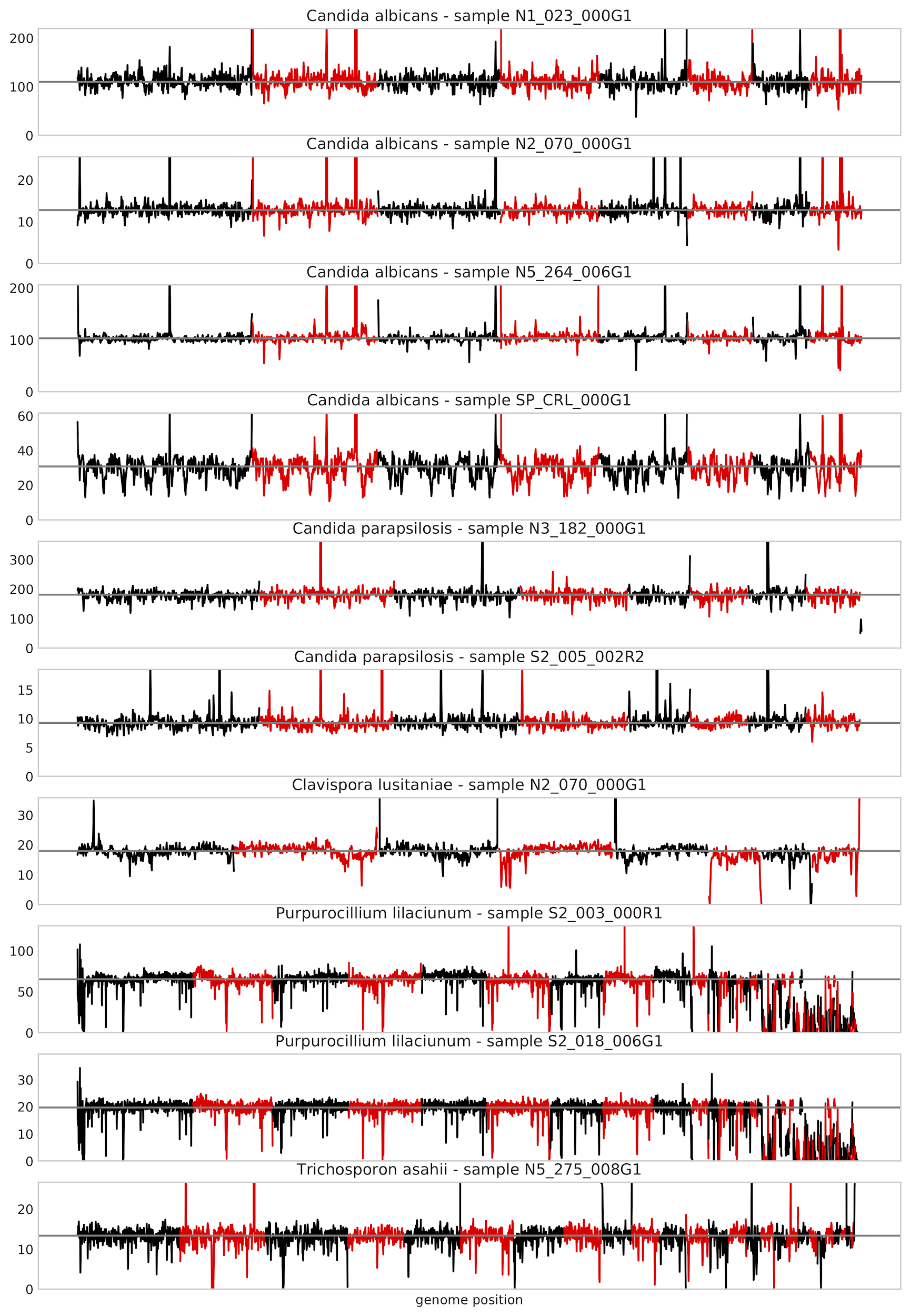
**Supplemental Figure S4.** Effect of coverage on variant frequency determination as assessed through simulation of metagenomic reads.

****

**Supplemental Figure S5.** Raw variant frequency graphs used to determine ploidy of all *de novo* assembled genomes.

****

**Supplemental Figure S6.** Determination of aneuploidy for all *de novo* assembled genomes based on scaffold coverage. The coverage of each 10kb window of each scaffold is shown. Scaffolds are ordered from largest to smallest, and rotate between red and black colors. No large portions of chromosomes were detected as having a multiple of 1/2x the coverage of the genome average as would be expected from a diploid genome. .



**Supplemental Figure S7.** Alternative mapping-based determination of aneuploidy for genomes with high quality reference genomes. No large portions of chromosomes were detected as having a multiple of 1/2x the coverage of the genome average as would be expected from a diploid genome.

**Supplemental Table S1**. Sequencing metadata for all infant and room metagenomic samples

**Supplemental Table S2.** Detailed information about genome assemblies

**Supplemental Table S3**. Mapping-based abundance of eukaryote genomes in all samples

**Supplemental Table S4**. Metadata for statistical associations

**Supplemental Table S5.** Statistical associations of samples containing eukaryotes with metadata