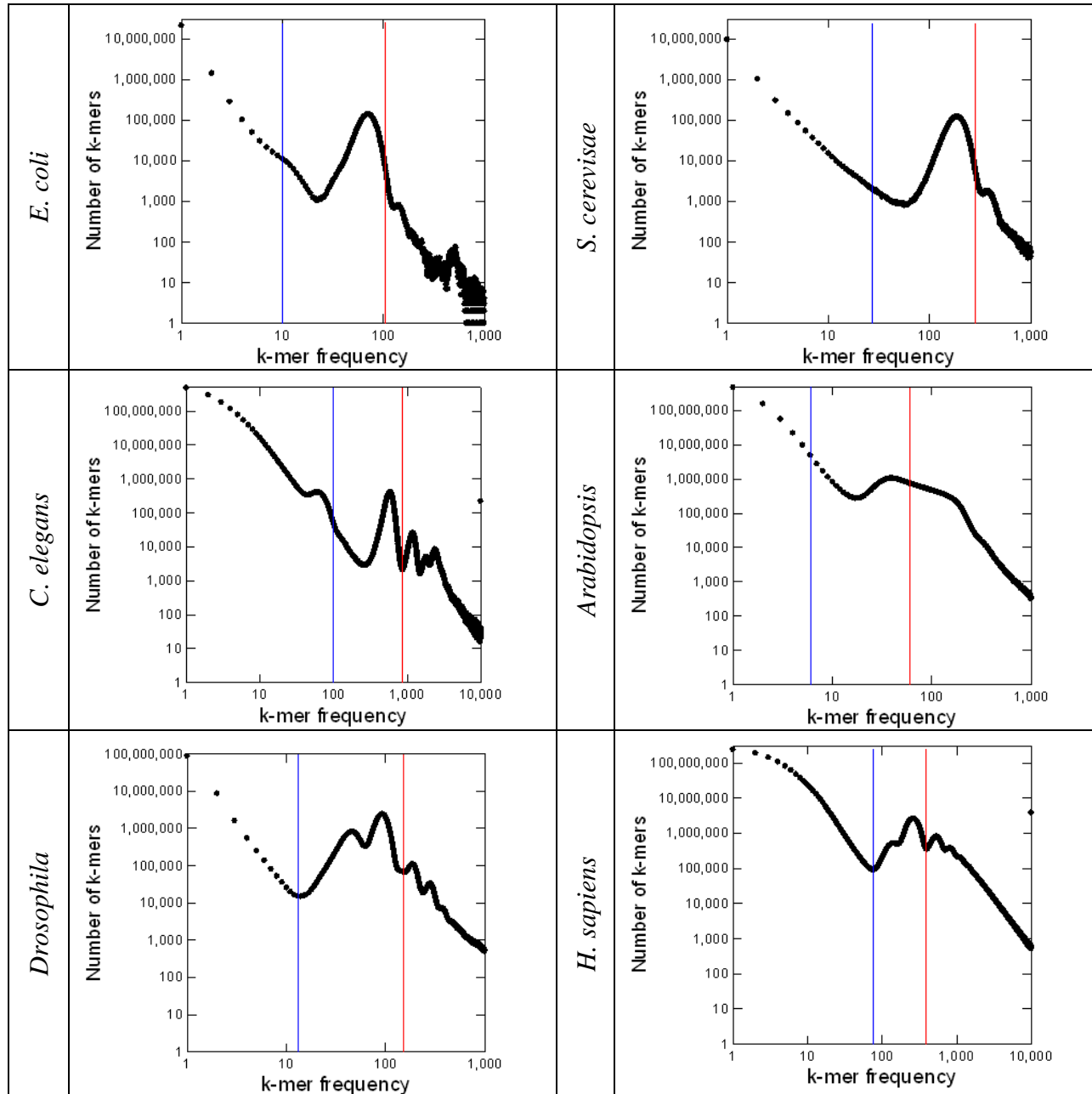


Supplemental Fig. S1



Supplemental Figure S1. Illumina  $k$ -mer frequency cut-offs. The blue line is the low-frequency cut-off, intended to remove error  $k$ -mers. The red line is the high-frequency cut-off, intended to remove  $k$ -mers that have more than one copy in the genome. Actual values are shown in Supplemental Table S2. In the *Drosophila* and *H.sapiens* data, the first peak is the X and Y single-copy  $k$ -mers, and the second peak is the autosomal ones (Supplemental Fig. S5).