

**Cell Systems, Volume 7**

**Supplemental Information**

**Mantis: A Fast, Small, and Exact**

**Large-Scale Sequence-Search Index**

**Prashant Pandey, Fatemeh Almodaresi, Michael A. Bender, Michael Ferdman, Rob Johnson, and Rob Patro**

## Table S1, related to Table 1, Table 2, and Table 3 – subsection Experimental Procedure

Min size	Max size	Cutoff
0	$\leq 300MB$	1
$> 300MB$	$\leq 500MB$	3
$> 500MB$	$\leq 1GB$	10
$> 1GB$	$\leq 3GB$	20
$> 3GB$	$\infty$	50

Table S1: Minimum number of times a k-mer must appear in an experiment in order to be counted as abundantly represented in that experiment (taken from Solomon & Kingsford (2016)).

## Figure S1, related to STAR Methods – subsection sampling color classes based on abundance

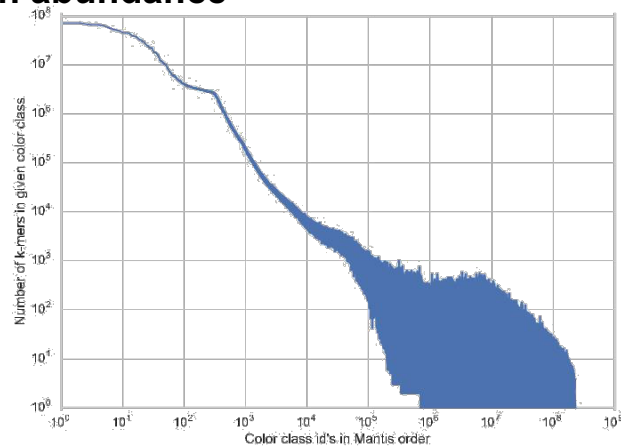


Figure S1: The distribution of the number of k-mers in a color class and ID assigned to the color class. The color class with the most number of k-mers gets the smallest ID. And the color class with least number of k-mers gets the largest ID. This distribu