



Supplementary Figure 1. Impact of the k-mer size on the performances of Simka. In each test above, Simka computed all the distances using the first 1M paired reads of 20 HMP samples on a single core. When k increases, the disk usage increases sub-linearly whereas the computation time and the memory usage stay constant.