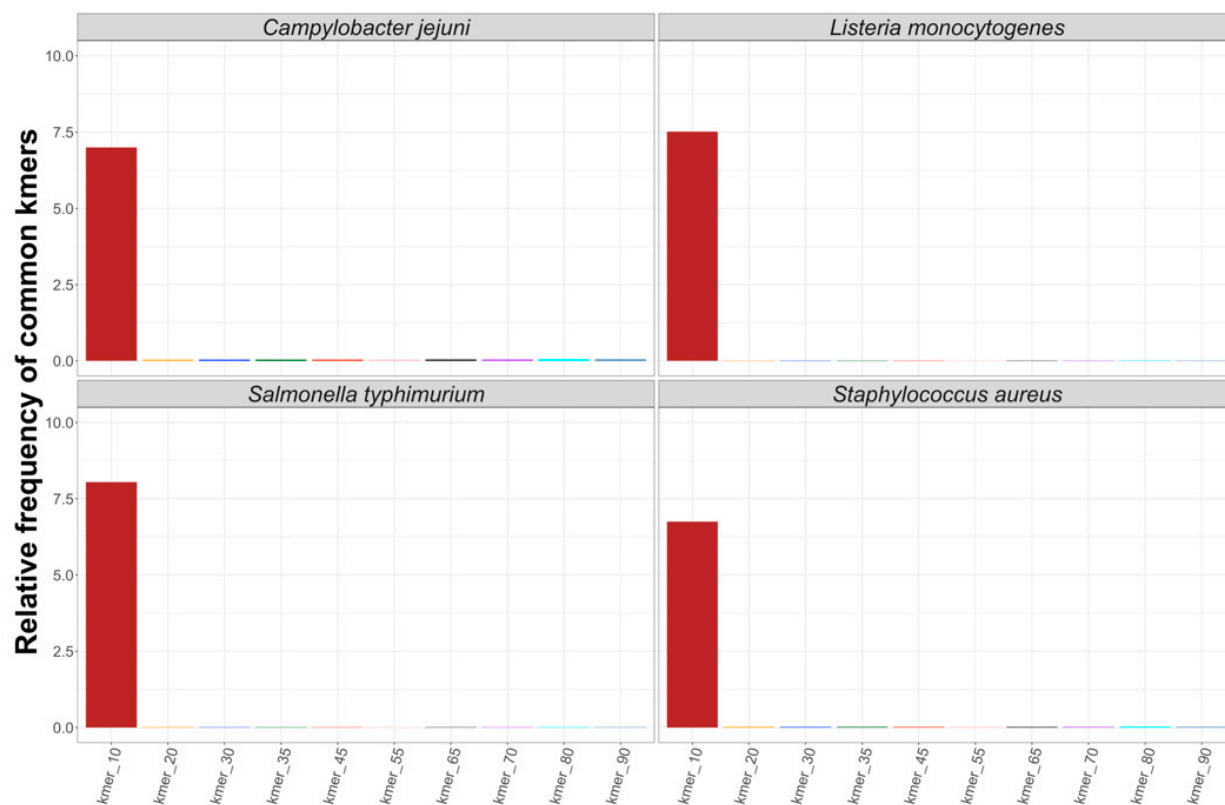


A.



B.

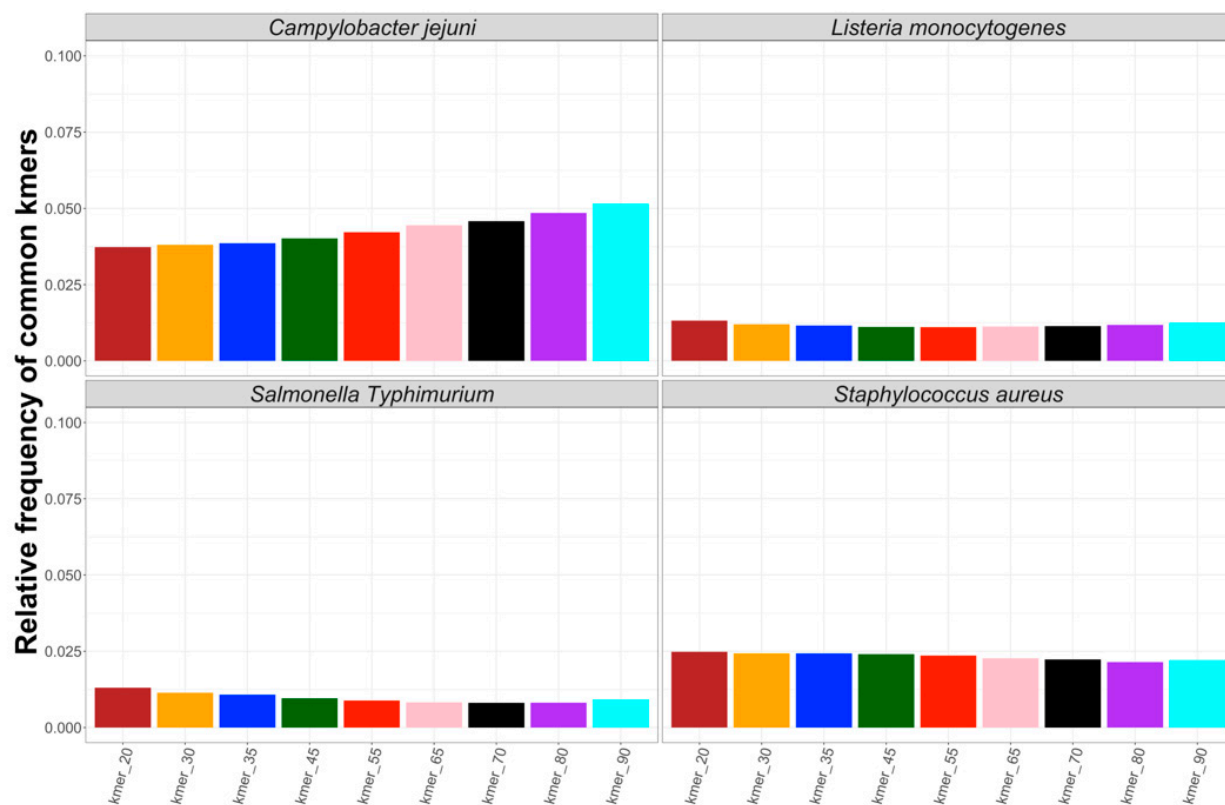


Figure S19. Relative frequency of common k-mers found in the raw reads and the stringMLST database.

Random 100 raw paired-end reads from the initial *C. jejuni*, *L. monocytogenes*, *S. aureus* and *S. Typhimurium* (major representative zoonotic serovar of *S. enterica*) datasets were selected, and DSK was used to count the occurrence of k-mers of lengths 10, 20, 30, 35, 45, 55, 65, 70, 80 and 90 respectively in the raw reads. (A) Next, for each database created with stringMLST, a file with the k-mer frequency for the used ST scheme was generated. Using the k-mers generated from the raw reads and the stringMLST database, a relative frequency of the common k-mers was calculated. (B) Due to the large range of frequency-based values between k-mer 10 and the remaining k-mer lengths, the k-mer 10 data were excluded from the plot.