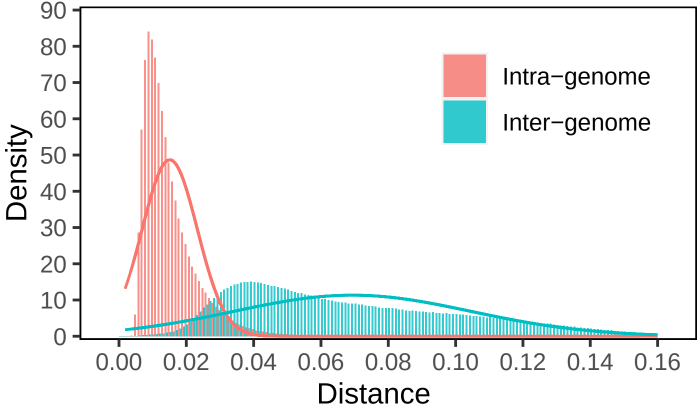

Supplementary Figure S12. Euclidean distance distributions between contigs abstracted from the same genome (intra-genome) and from two different genomes (inter-genome). The 3,023 bacterial genomes downloaded from NCBI were randomly fragmented into contigs (1 Kb - 100 Kb), then one million simulations were performed. Two histograms denote the intra- and inter- distances, respectively. Two curves represent Gaussian densities parametrized by the mean and variance of the intra- and inter- distances, respectively.