

**Figure S6.** Comparison of the distribution of some physicochemical properties in the top 1000 k-mers and in the full list of k-mers for height = 4.