



Supplementary Fig. S11. Distribution of the number of samples of a given SP type within each organism group. The partitions of the dataset are used for the nested cross-validation. The bold text in the x-axis represents the SP types with limited data.