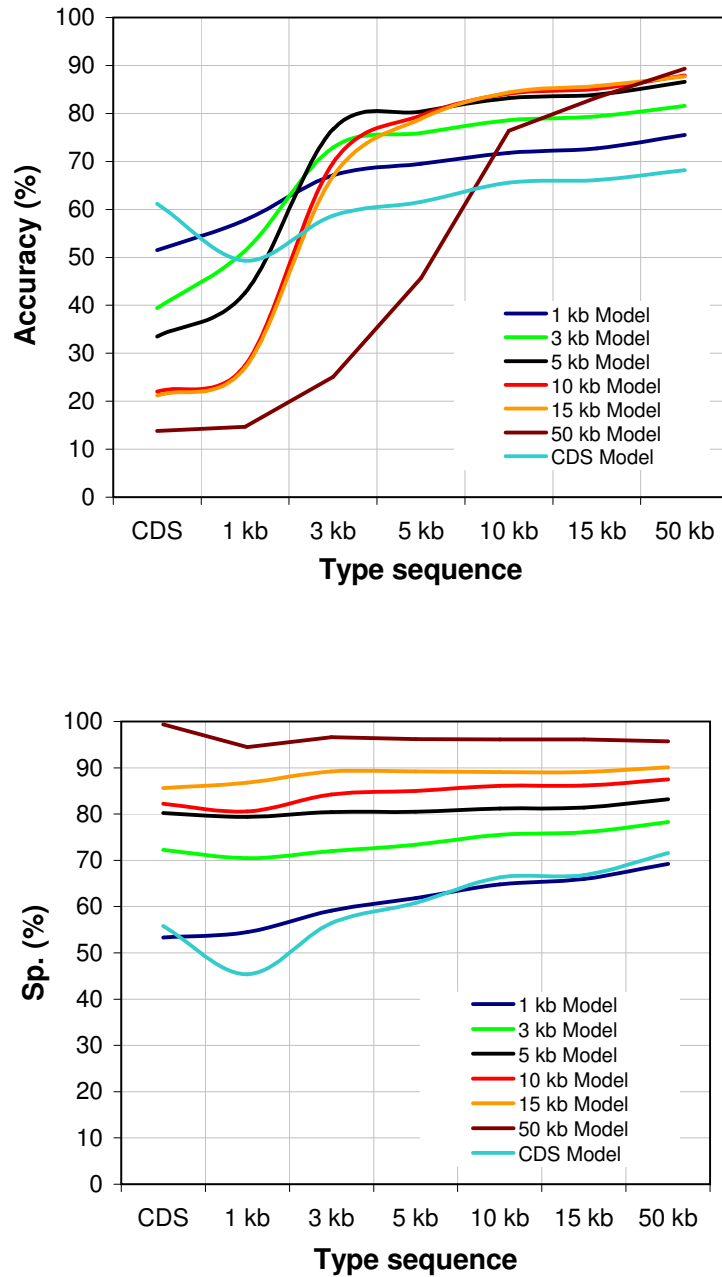Supplementary Figure 1. Assignment accuracy for differently sized genomic fragments (GF) and coding sequences (CDS) from unknown organisms at the level of the class. At the class level, 22 phylogenetic clades are included in the phylogenetic model. 7 class-level classifiers were trained with tetranucleotide patterns (w4, l4) derived from genomic fragments of length 1, 3, 5, 10, 15 and 50 kb as well as from CDSs. For each classifier, the accuracy (percentage of correctly assigned fragments), and specificity of assignments (*Sp*.) with fragments of different lengths is shown for organisms that are unknown to the model (of which no sequences were used in training).