

Supplementary Table 4

**Supplementary Table 4:** Phylogenetic classification accuracy of *PhyloPythia* for genomic fragments of known organisms at different taxonomic ranks. For every taxonomic level, the sensitivity (*Sn.*) and specificity (*Sp.*), averaged over all clades where 300+ fragments were available for classification) of *de novo* sequence classification is displayed. The specificity for 50 kb fragments at the genus level is undetermined (n.a.), due to too little available evaluation data.  $\Delta Sn.$  and  $\Delta Sp.$  give the improvement of classification accuracy compared to that for unknown organisms.

Rank	GF	<i>Sn.</i> (%)	<i>Sp.</i> (%)	$\Delta Sn.$	$\Delta Sp.$
<b>Domain</b>	1 kb	57.71	88.68	0.03	7.39
	3 kb	92.35	90.7	20.31	-0.6
	5 kb	96.03	98.28	26.88	1.36
	10 kb	98.15	98.03	11.74	5.96
	15 kb	97.59	99.72	20.73	1.28
	50 kb	99.8	99.95	10.61	2.03
<b>Phylum</b>	1 kb	50.32	92.11	9.77	12.64
	3 kb	94.75	94.84	19.07	12.89
	5 kb	96.53	94.73	17.25	12.35
	10 kb	97.93	97.22	17.91	11.77
	15 kb	98.77	98.61	19.3	10.92
	50 kb	97.99	98.82	17.42	11.03
<b>Class</b>	1 kb	30.76	93.21	0.07	10.17
	3 kb	94.64	94.29	15.4	10.87
	5 kb	96.93	93.75	12.96	11.76
	10 kb	97.65	96.19	10.67	14.45
	15 kb	98.9	98.72	11.31	17.09
	50 kb	99.53	98.45	11.69	15.61
<b>Order</b>	1 kb	25.12	95.04	18.73	-1.07
	3 kb	94.18	93.89	58.23	3.91
	5 kb	96.93	92.75	36.56	2.64
	10 kb	98.5	96.79	18.97	8.87
	15 kb	98.66	96.7	16.34	10.27
	50 kb	99.66	98.61	16.34	14.46
<b>Genus</b>	1 kb	7.11	96.72	2.69	8.84
	3 kb	69.16	92.5	30.54	1.4
	5 kb	95.43	89.18	32.05	1.57
	10 kb	98.07	83.68	16.98	0.02
	15 kb	96.63	87.82	9.11	5.22
	50 kb	99.52	n.a.	7.29	n.a.