Supplementary Table 1. W_n parameter search for the sequence composition space with the highest classification accuracy. Accuracy is shown for genomic fragments (upper part) and coding sequences (lower part) of 340 unknown organisms (to the classifier) at different phylogenetic levels.The composition space is defined by the word length *w*, the number of literal characters *l*, and the step size *s*. *Acc.* denotes the percentage of correctly assignments for all tested fragments, *Sn.* and *Sp.* are the sensitivity and specificity of the classification. Assignments were performed without the post-processing step in classification.

Rank	Data type	Wn	*Acc.* (%)	*Sn.* (%)	*Sp.* (%)
Genus	15 kb	w2,l2,s1	38.3	37	25.5
Genus	15 kb	w3,l2,s1	41.7	52.1	33.1
Genus	15 kb	w3,l3,s1	84.3	76.6	81.9
Genus	15 kb	w4,l4,s1	88.1	86.3	82.9
Genus	15 kb	w6,l4,s1	87.2	87.9	80.8
Genus	15 kb	**w5,l5,s1**	**87.6**	**89.1**	**80.4**
Genus	15 kb	w6,l5,s1	86.6	88.4	79.5
Genus	15 kb	w6,l6,s1	86.6	88.4	79.5
Order	15 kb	w2,l2,s1	43.3	33.1	62.7
Order	15 kb	w3,l2,s1	61.3	54.3	81
Order	15 kb	w3,l3,s1	77.1	70	84
Order	15 kb	w4,l4,s1	84.2	80.2	85.5
Order	15 kb	w6,l4,s1	84.9	81.1	83.9
Order	15 kb	**w5,l5,s1**	**85.1**	**81.4**	**85.5**
Order	15 kb	w6,l5,s1	84.9	81.1	83.9
Order	15 kb	w6,l6,s1	84.4	81	82.9
Class	15 kb	w2,l2,s1	43.2	39.6	68.5
Class	15 kb	w3,l2,s1	65.4	62.2	81.8
Class	15 kb	w3,l3,s1	77.9	74	88.1
Class	15 kb	w4,l4,s1	85.7	82.8	89.1
Class	15 kb	w6,l4,s1	87	83.8	88.8
Class	15 kb	**w5,l5,s1**	**87.1**	**84.4**	**88.9**
Class	15 kb	w6,l5,s1	87.2	83.8	88.5
Class	15 kb	w6,l6,s1	86.7	83.7	87.3
Phylum	15 kb	w2,l2,s1	43.1	37.2	34
Phylum	15 kb	w3,l2,s1	53.3	48.9	44.4
Phylum	15 kb	w3,l3,s1	59.6	56	51
Phylum	15 kb	w4,l4,s1	85.7	78.5	86.2
Phylum	15 kb	**w6,l4,s1**	**87.5**	**79.6**	**87.4**
Phylum	15 kb	w5,l5,s1	87.2	79.4	87.8
Phylum	15 kb	w6,l5,s1	87.4	79.1	87.5
Phylum	15 kb	w6,l6,s1	87	78.6	87.8
Domain	15 kb	w2,l2,s1	74.5	56.5	66.5
Domain	15 kb	w3,l2,s1	87.7	69	87.4
Domain	15 kb	w3,l3,s1	91	76.2	92.3
Domain	15 kb	w4,l4,s1	94.7	84.5	93.3
Domain	15 kb	w6,l4,s1	95.4	86.7	91.9
Domain	15 kb	w5,l5,s1	95.1	87.3	92.5
Domain	15 kb	w6,l5,s1	95.8	88.8	92.3
Domain	15 kb	**w6,l6,s1**	**95.7**	**90.1**	**92.5**

Supplementary Table 1

Rank	Data type	Wn	*Acc.* (%)	*Sn.* (%)	*Sp.* (%)
Genus	CDS	w2,l2,s1	*43.1*	*38.1*	*28.7*
Genus	CDS	w2,l2,s3	26.9	28.2	15.6
Genus	CDS	w3,l2,s3	40	57.8	31.7
Genus	CDS	w3,l3,s1	49.7	67.4	39.4
Genus	CDS	w3,l3,s3	49	67.1	38.5
Genus	CDS	w4,l4,s1	50.7	74.2	41.5
Genus	CDS	w4,l4,s3	51.7	75.3	42.4
Genus	CDS	w6,l4,s3	57	80.9	47.1
Genus	CDS	w5,l5,s1	52.7	76.5	43.4
Genus	CDS	w5,l5,s3	55.1	78	45.3
Genus	CDS	w6,l5,s3	57.1	81.2	47.2
Genus	CDS	w6,l6,s1	54.5	77.6	44.7
Genus	CDS	**w6,l6,s3**	**58.2**	**80**	**47.9**
Class	CDS	w2,l2,s1	34.3	36.2	35.7
Class	CDS	w2,l2,s3	19.6	27.5	26.2
Class	CDS	w3,l2,s3	40.2	50	40.6
Class	CDS	w3,l3,s1	51.1	57.5	49.2
Class	CDS	w3,l3,s3	53.9	60.9	51.3
Class	CDS	w4,l4,s1	57.1	65.3	52.5
Class	CDS	w4,l4,s3	58.3	67.9	54.5
Class	CDS	**w6,l4,s3**	**63.3**	**72.7**	**59.2**
Class	CDS	w5,l5,s1	57.6	66.9	53
Class	CDS	w5,l5,s3	59.4	69.8	56.6
Class	CDS	w6,l5,s3	62.3	71.9	58.5
Class	CDS	w6,l6,s1	57.8	67.7	54.2
Class	CDS	w6,l6,s3	60.6	70.2	57.5
Phylum	CDS	w2,l2,s1	40.2	37.4	35.5
Phylum	CDS	w2,l2,s3	26.3	37.1	30.3
Phylum	CDS	w3,l2,s3	42.4	53.9	40.5
Phylum	CDS	w3,l3,s1	57.2	57	49.1
Phylum	CDS	w3,l3,s3	58.1	58.4	50.8
Phylum	CDS	w4,l4,s1	60.2	63.2	50.8
Phylum	CDS	w4,l4,s3	58.9	65.1	50.8
Phylum	CDS	**w6,l4,s3**	**63**	**69**	**54.6**
Phylum	CDS	w5,l5,s1	57.5	63.3	48.9
Phylum	CDS	w5,l5,s3	57.7	66	51.4
Phylum	CDS	w6,l5,s3	61.8	68.2	54.3
Phylum	CDS	w6,l6,s1	56.6	63.5	49.2
Phylum	CDS	w6,l6,s3	59.9	66.6	52.7
Domain	CDS	w2,l2,s1	72.3	56.1	57.9
Domain	CDS	w2,l2,s3	61.8	61.5	54.7
Domain	CDS	w3,l2,s3	76.8	71.6	65.7
Domain	CDS	w3,l3,s1	76.6	69.4	62.9
Domain	CDS	w3,l3,s3	82.9	78.5	71.7
Domain	CDS	w4,l4,s1	83.4	74.2	71.8
Domain	CDS	w4,l4,s3	86	80.4	75
Domain	CDS	**w6,l4,s3**	**86.6**	**82.2**	**75.5**
Domain	CDS	w5,l5,s1	80	75.6	66.3
Domain	CDS	w5,l5,s3	84.1	79.7	72.1
Domain	CDS	w6,l5,s3	81.5	80	69.5
Domain	CDS	w6,l6,s1	n.d.	n.d.	n.d.
Domain	CDS	w6,l6,s3	84.4	81.1	74.1