

Table 1 Prediction accuracy evaluation of protist protein subcellular locations

	TP	FP	TN	FN	Sn	Sp	MCC
(a) Mitochondrial proteins							
TargetP	136	89	1823	359	27.5	95.3	0.32
WoLF PSORT	278	133	1779	217	56.2	93.0	0.53
TargetP AND WoLF PSORT	118	35	1877	377	23.8	98.2	0.36
TargetP OR WoLF PSORT	296	188	1724	199	59.8	90.2	0.50
(b) Secreted proteins							
Secreted	99	50	2211	47	67.8	97.8	0.65
S + HLS	130	85	2176	16	89.0	96.2	0.71
S + HLS + LS	137	121	2140	9	93.8	94.6	0.68
S+ HLS + LS + WLS	138	280	1981	8	94.5	87.6	0.52
(c) Other subcellular locations							
Cytoplasm	322	167	1714	204	61.2	91.1	0.54
Cytoskeleton	62	13	2180	152	29.0	99.4	0.46
ER	12	26	2254	115	9.4	98.9	0.15
Golgi	0	4	2345	58	0.0	99.8	0.00
Lysosome	0	0	2379	28	0.0	100.0	
Nucleus	466	514	1348	79	85.5	72.4	0.49
Peroxisome	1	2	2381	23	4.2	99.9	0.11
Plasma membrane	18	149	2046	194	8.5	93.2	0.02
Vacuole	0	0	2375	32	0.0	100.0	

Note: TP: true positives; FP: false positives; TN: true negatives; FN: false negatives. Sn: sensitivity; Sp: specificity; MCC: Matthews correlation coefficient. Secreted: predicted by 4 predictors; HLS: highly likely secreted, predicted by 3 out of 4 predictors; LS: likely secreted, predicted by 2 out of 4 predictors; WLS: weakly likely secreted, predicted by 1 out of 4 predictors