

Organism Name	Domain	Phylum	Genome Size (Mb)	G+C Content	Accession
<i>Pyrobaculum aerophilum</i> str. IM2	Archaea	Crenarchaeota	2.22243	51.36252	AE009441.1
<i>Pyrobaculum arsenaticum</i> DSM 13514	Archaea	Crenarchaeota	2.121076	55.05088	CP000660.1
<i>Pyrobaculum calidifontis</i> JCM 11548	Archaea	Crenarchaeota	2.009313	57.15083	CP000561.1
<i>Pyrobaculum islandicum</i> DSM 4184	Archaea	Crenarchaeota	1.826402	49.59905	CP000504.1
<i>Sulfolobus acidocaldarius</i> DSM 639	Archaea	Crenarchaeota	2.225959	36.70759	CP000077.1
<i>Sulfolobus solfataricus</i> P2	Archaea	Crenarchaeota	2.992245	35.78718	AE006641.1
<i>Sulfolobus tokodaii</i> str. 7	Archaea	Crenarchaeota	2.694756	32.78672	BA000023.2
<i>Methanococcus aeolicus</i> Nankai-3	Archaea	Euryarchaeota	1.5695	30.03746	CP000743.1
<i>Methanococcus maripaludis</i> C5	Archaea	Euryarchaeota	1.789046	32.97595	CP000609.1
<i>Methanococcus maripaludis</i> C6	Archaea	Euryarchaeota	1.789046	32.97595	CP000609.1
<i>Methanococcus maripaludis</i> C7	Archaea	Euryarchaeota	1.789046	32.97595	CP000609.1
<i>Methanococcus maripaludis</i> S2	Archaea	Euryarchaeota	1.661137	33.10034	BX950229.1
<i>Methanococcus vannielii</i> SB	Archaea	Euryarchaeota	1.720048	31.32831	CP000742.1
<i>Methanosarcina acetivorans</i> C2A	Archaea	Euryarchaeota	5.751492	42.68089	AE010299.1
<i>Methanosarcina barkeri</i> str. Fusaro	Archaea	Euryarchaeota	5.751492	42.68089	AE010299.1
<i>Methanosarcina mazei</i> Go1	Archaea	Euryarchaeota	4.096345	41.48491	AE008384.1
<i>Pyrococcus abyssi</i> GE5	Archaea	Euryarchaeota	1.768562	44.71141	AL096836.1
<i>Pyrococcus furiosus</i> DSM 3638	Archaea	Euryarchaeota	1.908256	40.7699	AE009950.1
<i>Pyrococcus horikoshii</i> OT3	Archaea	Euryarchaeota	1.738505	41.87897	BA000001.2
<i>Borrelia afzelii</i> PKo	Bacteria	Spirochaetes	1.404288	27.90168	CP002933.1
<i>Borrelia burgdorferi</i> B31	Bacteria	Spirochaetes	1.521208	28.17971	AE000783.1
<i>Borrelia duttonii</i> Ly	Bacteria	Spirochaetes	1.574881	28.01158	CP000976.1
<i>Borrelia garinii</i> PBi	Bacteria	Spirochaetes	0.986914	28.12697	CP000013.1
<i>Borrelia hermsii</i> DAH	Bacteria	Spirochaetes	1.498568	30.25355	CP073136.1
<i>Borrelia recurrentis</i> A1	Bacteria	Spirochaetes	1.242163	27.52553	CP000993.1
<i>Borrelia turicatae</i> 91E135	Bacteria	Spirochaetes	1.564967	29.3819	CP073176.1
<i>Leptospira biflexa</i> serovar Patoc strain Patoc 1 (Ames)	Bacteria	Spirochaetes	3.956089	38.89164	CP000777.1
<i>Leptospira biflexa</i> serovar Patoc strain Patoc 1 (Paris)	Bacteria	Spirochaetes	3.951448	38.89465	CP000786.1
<i>Leptospira borgpetersenii</i> serovar Hardjo-bovis JB197	Bacteria	Spirochaetes	3.876235	40.24403	CP000350.1
<i>Leptospira borgpetersenii</i> serovar Hardjo-bovis L550	Bacteria	Spirochaetes	3.931782	40.22911	CP000348.1
<i>Leptospira interrogans</i> serovar Copenhageni str. Fiocruz L1-130	Bacteria	Spirochaetes	4.627366	35.04434	AE016823.1
<i>Leptospira interrogans</i> serovar Lai str. 56601	Bacteria	Spirochaetes	4.698134	35.02493	AE010300.2
<i>Treponema denticola</i> ATCC 35405	Bacteria	Spirochaetes	2.843201	37.87376	AE017226.1
<i>Treponema pallidum</i> subsp. pallidum SS14	Bacteria	Spirochaetes	1.139457	52.76031	CP000805.1
<i>Treponema pallidum</i> subsp. pallidum str. Nichols	Bacteria	Spirochaetes	1.139633	52.78831	CP004010.2
<i>Mycoplasma agalactiae</i> PG2	Bacteria	Tenericutes	0.877438	29.70706	CU179680.1
<i>Mycoplasma arthritidis</i> 158L3-1	Bacteria	Tenericutes	0.820453	30.7095	CP001047.1

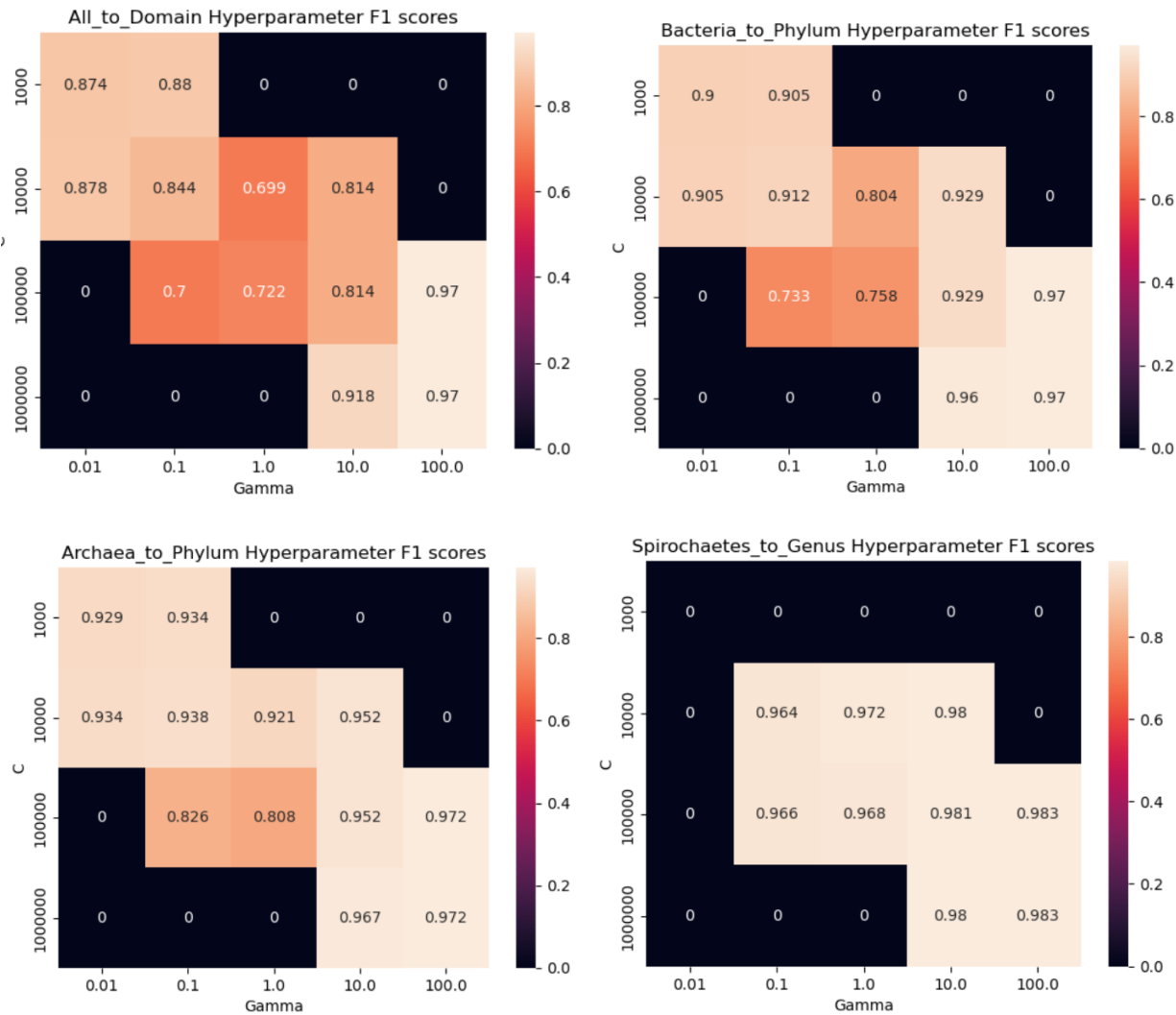
Mycoplasma capricolum subsp. capricolum ATCC 27343	Bacteria	Tenericutes	1.010023	23.77045	CP000123.1
Mycoplasma gallisepticum R	Bacteria	Tenericutes	1.0128	31.46989	AE015450.2
Mycoplasma genitalium G37	Bacteria	Tenericutes	0.580076	31.68912	L43967.2
Mycoplasma hyopneumoniae 232	Bacteria	Tenericutes	0.892758	28.56116	AE017332.1
Mycoplasma hyopneumoniae 7448	Bacteria	Tenericutes	0.920079	28.48886	AE017244.1
Mycoplasma hyopneumoniae J	Bacteria	Tenericutes	0.897405	28.52224	AE017243.1
Mycoplasma mobile 163K	Bacteria	Tenericutes	0.777079	24.95126	AE017308.1
Mycoplasma mycoides subsp. mycoides SC str. PG1	Bacteria	Tenericutes	1.211703	23.96561	BX293980.2
Mycoplasma penetrans HF-2	Bacteria	Tenericutes	1.358633	25.71754	BA000026.2
Mycoplasma pneumoniae M129	Bacteria	Tenericutes	0.816394	40.00801	U00089.2
Mycoplasma pulmonis UAB CTIP	Bacteria	Tenericutes	0.963879	26.63654	AL445566.1
Mycoplasma synoviae 53	Bacteria	Tenericutes	0.799476	28.49516	AE017245.1
Ureaplasma parvum serovar 3 str. ATCC 27815	Bacteria	Tenericutes	0.751679	25.49652	CP000942.1
Ureaplasma parvum serovar 3 str. ATCC 700970	Bacteria	Tenericutes	0.751719	25.49969	AF222894.1
Ureaplasma urealyticum serovar 10 str. ATCC 33699	Bacteria	Tenericutes	0.874478	25.77423	CP001184.1

Table S1. Prokaryotic genomes used during training.

Organism Name	Domain	Phylum	Genome Size (Mb)	G+C Content	Accession
Pyrobaculum oguniense	Archaea	Crenarchaeota	2.45292	55.051	CP003316.1
Pyrobaculum neutrophilum	Archaea	Crenarchaeota	1.769823	59.91062	CP001014.1
Pyrobaculum ferrireducens	Archaea	Crenarchaeota	2.467972	56.96681	CP003098.1
Sulfolobus islandicus	Archaea	Crenarchaeota	2.655201	35.14819	CP002426.1
Methanococcus voltae	Archaea	Euryarchaeota	1.857959	28.51931	JANUCR01000001.1
Methanosarcina siciliae	Archaea	Euryarchaeota	5.017558	42.94041	CP009506.1
Methanosarcina thermophila	Archaea	Euryarchaeota	3.100602	41.27053	FPA001000001.1
Methanosarcina horonobensis	Archaea	Euryarchaeota	5.018607	41.31105	CP009516.1
Pyrococcus yayanosii	Archaea	Euryarchaeota	1.716818	51.63751	CP002779.1
Pyrococcus kukulkanii	Archaea	Euryarchaeota	1.977126	44.62209	CP010835.1
Leptospira kirschneri	Bacteria	Spirochaetes	4.414224	36.04516	LSSQ01000001.1
Leptospira santarosai	Bacteria	Spirochaetes	3.835045	41.79422	APGN01000001.1
Leptospira weilii	Bacteria	Spirochaetes	4.370034	40.86421	AFLV02000084.1
Treponema pedis	Bacteria	Spirochaetes	3.077465	36.83636	CP061839.1
Treponema vincentii	Bacteria	Spirochaetes	2.594485	45.67658	CP051197.1
Treponema primitia	Bacteria	Spirochaetes	4.059867	50.83753	CP001843.1
Borrelia parkeri	Bacteria	Spirochaetes	1.030684	28.88208	CP007022.1
Borrelia anserina	Bacteria	Spirochaetes	1.04269	29.36769	CP013704.1
Borrelia miyamotoi	Bacteria	Spirochaetes	0.907293	28.71895	CP010308.1
Mycoplasma leachii	Bacteria	Tenericutes	1.008951	23.75051	CP002108.1
Mycoplasma yeatsii	Bacteria	Tenericutes	0.895051	25.73563	CP007520.1

Mycoplasma putrefaciens	Bacteria	Tenericutes	0.859996	26.96175	CP004357.1
Ureaplasma diversum	Bacteria	Tenericutes	1.011202	28.08598	JFDP01000001.1
Ureaplasma canigenitalium	Bacteria	Tenericutes	0.835388	31.85442	KL544010.1
Ureaplasma miroungigenitalium	Bacteria	Tenericutes	0.728586	31.39451	JA0XHL010000009.1

Table S2. Prokaryotic genomes used during unknown organism prediction tests.



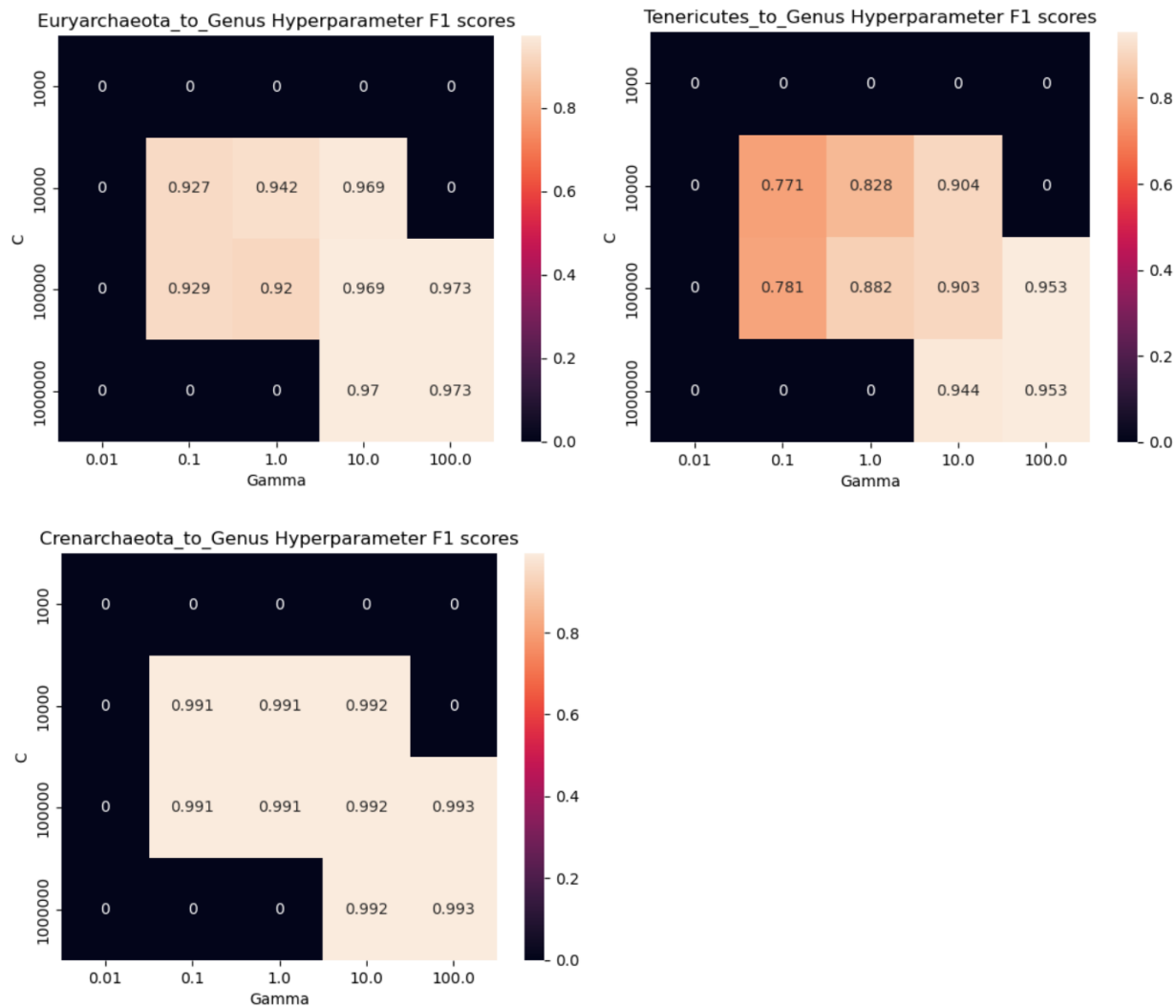


Figure S1. GridSearchCV hyperparameter tuning results for 7 support vector machine classifiers.