

I. ACCESSION NUMBERS

Organism	Accession Number in GeneBank FTP
<i>Mycoplasma pulmonis</i>	NC_002771
<i>Mycobacterium tuberculosis</i> CDC1551	NC_002755
<i>Treponema pallidum</i>	NC_000919
<i>Helicobacter pylori</i> 26695	NC_000915
<i>Pseudomonas aeruginosa</i>	NC_002516
<i>Borrelia burgdorferi</i>	NC_001318
<i>Haemophilus influenzae</i> Rd	NC_000907
<i>Salmonella enterica</i>	NC_003198
<i>Staphylococcus aureus</i> strain Mu50	NC_002758
<i>Escherichia coli</i>	NC_000913
<i>Bacillus subtilis</i>	NC_000964
<i>Lactococcus lactis</i> subsp. lactis	NC_002662
<i>Saccharomyces cerevisiae</i>	NC_001133 - NC_001148
<i>Caenorhabditis elegans</i>	chrI - chrX
<i>Drosophila melanogaster</i>	AE002566, AE002575, AE002546, AE002593 AE002602, AE002620, AE002629, AE002638 AE002647, AE002681, AE002690, AE002699 AE002708, AE002725, AE002769, AE002778 AE002787, AE002796, AE002804

Table 1. Full list of genomes considered in the paper with their accession numbers in GeneBank FTP. For *Saccharomyces cerevisiae*, the accession numbers go from NC_001133 to NC_001148; for *Caenorhabditis elegans*, the accession numbers cover all six chromosomes I, II, III, IV, V and X.

II. TABLES OF WEIGHTS $w_{i,j}$ COMPUTED FROM REFERENCE SET

ttt	1.000	ttc	0.048	ttg	0.062	tta	1.000	tct	0.638	tcc	0.043	tcg	0.043	tca	1.000
tgt	1.000	tgc	0.000	tgg	0.000	tga	1.000	tat	1.000	tac	0.069	tag	0.000	taa	0.000
ctt	0.123	ctc	0.031	ctg	0.000	cta	0.123	cct	1.000	ccc	0.000	cgc	0.000	cca	0.289
cgt	0.037	cgc	0.000	cgg	0.000	cga	0.000	cat	1.000	cac	0.125	cag	0.016	caa	1.000
gtt	0.778	gtc	0.074	gtg	0.074	gta	1.000	gct	1.000	gcc	0.023	gcg	0.023	gca	0.568
ggt	0.421	ggc	0.000	ggg	0.000	gga	1.000	gat	1.000	gac	0.248	gag	0.042	gaa	1.000
att	1.000	atc	0.140	atg	1.000	ata	0.488	act	1.000	acc	0.032	acg	0.000	aca	0.832
agt	0.809	agc	0.043	agg	0.000	aga	1.000	aat	1.000	aac	0.148	aag	0.062	aaa	1.000

Weights for *M. pulmonis*.

ttt	0.105	ttc	1.000	ttg	0.160	tta	0.007	tct	0.044	tcc	0.711	tcg	1.000	tca	0.048
tgt	0.127	tgc	1.000	tgg	1.000	tga	0.000	tat	0.162	tac	1.000	tag	0.000	taa	0.000
ctt	0.022	ctc	0.249	ctg	1.000	cta	0.027	cct	0.029	ccc	0.528	cgc	1.000	cca	0.060
cgt	0.150	cgc	1.000	cgg	0.756	cga	0.080	cat	0.170	cac	1.000	cag	1.000	caa	0.145
gtt	0.070	gtc	0.802	gtg	1.000	gta	0.029	gct	0.051	gcc	1.000	gcg	0.660	gca	0.075
ggt	0.218	ggc	1.000	ggg	0.251	gga	0.069	gat	0.142	gac	1.000	gag	1.000	gaa	0.351
att	0.062	atc	1.000	atg	1.000	ata	0.014	act	0.033	acc	1.000	acg	0.236	aca	0.030
agt	0.052	agc	0.693	agg	0.042	aga	0.004	aat	0.083	aac	1.000	aag	1.000	aaa	0.168

Weights for *M. tuberculosis*.

ttt	1.000	ttc	0.200	ttg	1.000	tta	0.134	tct	1.000	tcc	0.421	tcg	0.895	tca	0.526
tgt	1.000	tgc	0.417	tgg	1.000	tga	0.000	tat	1.000	tac	0.412	tag	0.000	taa	0.000
ctt	0.493	ctc	0.239	ctg	0.806	cta	0.075	cct	0.440	ccc	0.280	ccg	1.000	cca	0.080
cgt	1.000	cgc	0.868	cgg	0.974	cga	0.105	cat	0.632	cac	1.000	cag	1.000	caa	0.206
gtt	0.211	gtc	0.105	gtg	1.000	gta	0.179	gct	0.110	gcc	0.140	gcg	1.000	gca	0.200
ggg	0.536	ggc	0.161	ggg	1.000	gga	0.304	gat	1.000	gac	0.784	gag	1.000	gaa	0.589
att	1.000	atc	0.162	atg	1.000	ata	0.243	act	0.125	acc	0.156	acg	1.000	aca	0.250
agt	1.000	agc	0.737	agg	0.158	aga	0.026	aat	1.000	aac	0.737	aag	1.000	aaa	0.275

Weights for *T. pallidum*.

ttt	1.000	ttc	0.055	ttg	0.628	tta	1.000	tct	0.692	tcc	0.051	tcg	0.051	tca	0.179
tgt	0.125	tgc	1.000	tgg	1.000	tga	0.000	tat	1.000	tac	0.261	tag	0.000	taa	0.000
ctt	0.295	ctc	0.051	ctg	0.013	cta	0.038	cct	1.000	ccc	0.067	ccg	0.133	cca	0.467
cgt	0.333	cgc	0.800	cgg	0.000	cga	0.333	cat	1.000	cac	0.200	cag	0.016	caa	1.000
gtt	0.571	gtc	0.286	gtg	1.000	gta	0.171	gct	0.912	gcc	0.265	gcg	1.000	gca	0.235
ggg	0.227	ggc	0.545	ggg	1.000	gga	0.045	gat	1.000	gac	0.221	gag	0.267	gaa	1.000
att	1.000	atc	0.667	atg	1.000	ata	0.176	act	0.556	acc	1.000	acg	0.333	aca	0.333
agt	0.282	agc	1.000	agg	0.733	aga	1.000	aat	1.000	aac	0.704	aag	0.219	aaa	1.000

Weights for *H. pylori*.

ttt	0.001	ttc	1.000	ttg	0.022	tta	0.000	tct	0.000	tcc	0.934	tcg	0.577	tca	0.003
tgt	0.016	tgc	1.000	tgg	1.000	tga	0.000	tat	0.118	tac	1.000	tag	0.000	taa	0.000
ctt	0.004	ctc	0.288	ctg	1.000	cta	0.002	cct	0.011	ccc	0.269	ccg	1.000	cca	0.009
cgt	0.095	cgc	1.000	cgg	0.070	cga	0.001	cat	0.134	cac	1.000	cag	1.000	caa	0.093
gtt	0.047	gtc	1.000	gtg	0.979	gta	0.065	gct	0.034	gcc	1.000	gcg	0.244	gca	0.029
ggg	0.073	ggc	1.000	ggg	0.030	gga	0.007	gat	0.099	gac	1.000	gag	1.000	gaa	0.612
att	0.026	atc	1.000	atg	1.000	ata	0.001	act	0.018	acc	1.000	acg	0.031	aca	0.002
agt	0.019	agc	1.000	agg	0.013	aga	0.000	aat	0.043	aac	1.000	aag	1.000	aaa	0.075

Weights for *P. aeruginosa*

ttt	1.000	ttc	0.000	ttg	0.571	tta	1.000	tct	0.667	tcc	0.000	tcg	0.111	tca	0.444
tgt	1.000	tgc	0.091	tgg	1.000	tga	0.000	tat	1.000	tac	0.000	tag	0.000	taa	0.000
ctt	0.333	ctc	0.000	ctg	0.000	cta	0.000	cct	1.000	ccc	0.000	ccg	0.000	cca	0.500
cgt	0.000	cgc	0.000	cgg	0.000	cga	0.059	cat	1.000	cac	0.100	cag	0.000	caa	1.000
gtt	1.000	gtc	0.000	gtg	0.167	gta	0.375	gct	1.000	gcc	0.000	gcg	0.000	gca	1.000
ggg	1.000	ggc	0.500	ggg	0.333	gga	0.500	gat	1.000	gac	0.000	gag	0.333	gaa	1.000
att	1.000	atc	0.000	atg	1.000	ata	0.567	act	1.000	acc	0.000	acg	0.000	aca	0.167
agt	1.000	agc	0.111	agg	0.059	aga	1.000	aat	1.000	aac	0.000	aag	0.346	aaa	1.000

Weights for *B. burgdorferi*

ttt	0.483	ttc	1.000	ttg	0.049	tta	1.000	tct	1.000	tcc	0.000	tcg	0.000	tca	0.375
tgt	1.000	tgc	0.300	tgg	1.000	tga	0.000	tat	0.390	tac	1.000	tag	0.000	taa	0.000
ctt	0.174	ctc	0.000	ctg	0.005	cta	0.022	cct	0.312	ccc	0.000	ccg	0.000	cca	1.000
cgt	1.000	cgc	0.113	cgg	0.000	cga	0.000	cat	0.205	cac	1.000	cag	0.022	caa	1.000
gtt	1.000	gtc	0.033	gtg	0.197	gta	0.711	gct	0.574	gcc	0.010	gcg	0.206	gca	1.000
ggg	1.000	ggc	0.326	ggg	0.005	gga	0.016	gat	1.000	gac	0.348	gag	0.088	gaa	1.000
att	0.482	atc	1.000	atg	1.000	ata	0.000	act	1.000	acc	0.106	acg	0.080	aca	0.195
agt	0.135	agc	0.156	agg	0.000	aga	0.008	aat	0.484	aac	1.000	aag	0.026	aaa	1.000

Weights for *H. influenzae*.

ttt	0.371	ttc	1.000	ttg	0.560	tta	0.028	tct	0.400	tcc	0.000	tcg	0.000	tca	1.000
tgt	1.000	tgc	0.100	tgg	1.000	tga	0.000	tat	0.208	tac	1.000	tag	0.000	taa	0.000
ctt	1.000	ctc	0.188	ctg	0.000	cta	0.000	cct	0.468	ccc	0.000	ccg	0.008	cca	1.000
cgt	1.000	cgc	0.171	cgg	0.000	cga	0.000	cat	0.371	cac	1.000	cag	0.000	caa	1.000
gtt	1.000	gtc	0.060	gtg	0.038	gta	0.350	gct	1.000	gcc	0.037	gcg	0.066	gca	0.536
ggt	1.000	ggc	0.081	ggg	0.006	gga	0.281	gat	1.000	gac	0.994	gag	0.009	gaa	1.000
att	0.324	atc	1.000	atg	1.000	ata	0.000	act	1.000	acc	0.010	acg	0.000	aca	0.608
agt	0.024	agc	0.073	agg	0.000	aga	0.005	aat	0.295	aac	1.000	aag	0.025	aaa	1.000

Weights for *L. lactis*.

ttt	0.518	ttc	1.000	ttg	0.003	tta	1.000	tct	0.381	tcc	0.000	tcg	0.000	tca	1.000
tgt	1.000	tgc	0.500	tgg	1.000	tga	0.000	tat	0.676	tac	1.000	tag	0.000	taa	0.000
ctt	0.083	ctc	0.000	ctg	0.000	cta	0.020	cct	0.402	ccc	0.000	ccg	0.029	cca	1.000
cgt	1.000	cgc	0.068	cgg	0.000	cga	0.007	cat	0.881	cac	1.000	cag	0.009	caa	1.000
gtt	1.000	gtc	0.013	gtg	0.038	gta	0.843	gct	1.000	gcc	0.009	gcg	0.088	gca	0.912
ggt	1.000	ggc	0.120	ggg	0.003	gga	0.172	gat	1.000	gac	0.746	gag	0.066	gaa	1.000
att	0.843	atc	1.000	atg	1.000	ata	0.012	act	1.000	acc	0.006	acg	0.022	aca	0.696
agt	0.194	agc	0.209	agg	0.000	aga	0.162	aat	0.774	aac	1.000	aag	0.026	aaa	1.000

Weights for *S. aureus*.

ttt	0.292	ttc	1.000	ttg	0.017	tta	0.013	tct	1.000	tcc	0.753	tcg	0.030	tca	0.013
tgt	0.279	tgc	1.000	tgg	1.000	tga	0.000	tat	0.301	tac	1.000	tag	0.000	taa	0.000
ctt	0.030	ctc	0.030	ctg	1.000	cta	0.001	cct	0.099	ccc	0.003	ccg	1.000	cca	0.114
cgt	1.000	cgc	0.465	cgg	0.000	cga	0.002	cat	0.185	cac	1.000	cag	1.000	caa	0.109
gtt	1.000	gtc	0.154	gtg	0.366	gta	0.431	gct	1.000	gcc	0.261	gcg	0.649	gca	0.436
ggt	0.925	ggc	1.000	ggg	0.022	gga	0.009	gat	0.490	gac	1.000	gag	0.299	gaa	1.000
att	0.265	atc	1.000	atg	1.000	ata	0.000	act	0.617	acc	1.000	acg	0.093	aca	0.051
agt	0.021	agc	0.447	agg	0.000	aga	0.000	aat	0.101	aac	1.000	aag	0.223	aaa	1.000

Weights for *S. enterica*.

ttt	0.337	ttc	1.000	ttg	0.139	tta	0.377	tct	1.000	tcc	0.071	tcg	0.015	tca	0.301
tgt	0.529	tgc	1.000	tgg	1.000	tga	0.000	tat	0.261	tac	1.000	tag	0.000	taa	0.000
ctt	1.000	ctc	0.035	ctg	0.098	cta	0.104	cct	0.939	ccc	0.017	ccg	0.226	cca	1.000
cgt	1.000	cgc	0.535	cgg	0.000	cga	0.000	cat	0.583	cac	1.000	cag	0.183	caa	1.000
gtt	1.000	gtc	0.117	gtg	0.150	gta	0.585	gct	1.000	gcc	0.046	gcg	0.268	gca	0.528
ggt	1.000	ggc	0.767	ggg	0.022	gga	0.648	gat	1.000	gac	0.880	gag	0.208	gaa	1.000
att	0.378	atc	1.000	atg	1.000	ata	0.006	act	1.000	acc	0.010	acg	0.141	aca	0.932
agt	0.041	agc	0.265	agg	0.000	aga	0.066	aat	0.208	aac	1.000	aag	0.116	aaa	1.000

Weights for *B. subtilis*.

ttt	0.188	ttc	1.000	ttg	0.018	tta	0.012	tct	1.000	tcc	0.627	tcg	0.008	tca	0.025
tgt	0.277	tgc	1.000	tgg	1.000	tga	0.000	tat	0.240	tac	1.000	tag	0.000	taa	0.000
ctt	0.025	ctc	0.030	ctg	1.000	cta	0.003	cct	0.084	ccc	0.003	ccg	1.000	cca	0.151
cgt	1.000	cgc	0.419	cgg	0.000	cga	0.003	cat	0.244	cac	1.000	cag	1.000	caa	0.110
gtt	1.000	gtc	0.060	gtg	0.227	gta	0.504	gct	1.000	gcc	0.112	gcg	0.359	gca	0.515
ggt	1.000	ggc	0.694	ggg	0.012	gga	0.005	gat	0.396	gac	1.000	gag	0.182	gaa	1.000
att	0.201	atc	1.000	atg	1.000	ata	0.002	act	0.805	acc	1.000	acg	0.032	aca	0.037
agt	0.025	agc	0.477	agg	0.000	aga	0.000	aat	0.051	aac	1.000	aag	0.196	aaa	1.000

Weights for *E. coli*.

ttt	0.136	ttc	1.000	ttg	1.000	tta	0.124	tct	1.000	tcc	0.723	tcg	0.000	tca	0.015
tgt	1.000	tgc	0.066	tgg	1.000	tga	0.000	tat	0.067	tac	1.000	tag	0.000	taa	0.000
ctt	0.005	ctc	0.000	ctg	0.001	cta	0.036	cct	0.084	ccc	0.003	ccg	0.000	cca	1.000
cgt	0.141	cgc	0.000	cgg	0.000	cga	0.000	cat	0.199	cac	1.000	cag	0.000	caa	1.000
gtt	1.000	gtc	0.823	gtg	0.018	gta	0.000	gct	1.000	gcc	0.300	gcg	0.001	gca	0.004
ggg	1.000	ggc	0.021	ggg	0.003	gga	0.003	gat	0.621	gac	1.000	gag	0.012	gaa	1.000
att	0.747	atc	1.000	atg	1.000	ata	0.002	act	0.966	acc	1.000	acg	0.000	aca	0.012
agt	0.024	agc	0.031	agg	0.000	aga	1.000	aat	0.052	aac	1.000	aag	1.000	aaa	0.095

Weights for *S. cerevisiae*.

ttt	0.060	ttc	1.000	ttg	0.288	tta	0.004	tct	0.636	tcc	1.000	tcg	0.147	tca	0.193
tgt	0.187	tgc	1.000	tgg	1.000	tga	0.000	tat	0.152	tac	1.000	tag	0.000	taa	0.000
ctt	0.830	ctc	1.000	ctg	0.049	cta	0.004	cct	0.013	ccc	0.006	ccg	0.012	cca	1.000
cgt	1.000	cgc	0.645	cgg	0.005	cga	0.040	cat	0.355	cac	1.000	cag	0.367	caa	1.000
gtt	0.652	gtc	1.000	gtg	0.134	gta	0.045	gct	0.823	gcc	1.000	gcg	0.013	gca	0.122
ggg	0.057	ggc	0.021	ggg	0.006	gga	1.000	gat	0.819	gac	1.000	gag	1.000	gaa	0.402
att	0.285	atc	1.000	atg	1.000	ata	0.005	act	0.405	acc	1.000	acg	0.024	aca	0.065
agt	0.035	agc	0.190	agg	0.011	aga	0.488	aat	0.187	aac	1.000	aag	1.000	aaa	0.088

Weights for *C. elegans*.

ttt	0.056	ttc	1.000	ttg	0.130	tta	0.003	tct	0.111	tcc	1.000	tcg	0.413	tca	0.030
tgt	0.084	tgc	1.000	tgg	1.000	tga	0.000	tat	0.121	tac	1.000	tag	0.000	taa	0.000
ctt	0.054	ctc	0.236	ctg	1.000	cta	0.019	cct	0.095	ccc	1.000	ccg	0.157	cca	0.148
cgt	0.478	cgc	1.000	cgg	0.032	cga	0.024	cat	0.188	cac	1.000	cag	1.000	caa	0.079
gtt	0.255	gtc	0.734	gtg	1.000	gta	0.028	gct	0.233	gcc	1.000	gcg	0.069	gca	0.035
ggg	0.500	ggc	1.000	ggg	0.006	gga	0.359	gat	0.597	gac	1.000	gag	1.000	gaa	0.086
att	0.231	atc	1.000	atg	1.000	ata	0.009	act	0.112	acc	1.000	acg	0.104	aca	0.040
agt	0.045	agc	0.512	agg	0.080	aga	0.013	aat	0.127	aac	1.000	aag	1.000	aaa	0.050

Weights for *D. melanogaster*.

III. CORRELATION TABLE

	B.subtilis	B.burgdorferi	E.coli	H.influenzae	H.pylori
GC_CONT	-0.20259733	-0.053434618	0.3141856	0.02257051	-0.11091986
GC_CONT3	-0.45153016	-0.42543584	0.32064584	-0.44750726	-0.13217682
GC_SKEW	-0.043287214	0.7767645	-0.11297992	0.0520236	0.35810432
AT_SKEW	0.2889337	-0.5517935	0.1884775	0.1785525	0.13681133

	L. lactis	M. pulmonis	P. aeruginosa	S. enterica	S. aureus
GC_CONT	0.3606557	-0.29194045	0.33130258	0.36331645	0.12785803
GC_CONT3	-0.2208189	-0.8173068	0.83097416	0.4133271	-0.53375053
GC_SKEW	-0.20575686	0.18914002	-0.46513942	-0.10253788	0.021157095
AT_SKEW	0.0752821	0.24208333	0.2397369	0.16666141	0.29059884

	T. pallidum	M. tuberculosis	S. cerevisiae	C. elegans	D. melanogaster
GC_CONT	0.074771784	0.44744366	0.19829214	0.5605825	0.62064445
GC_CONT3	0.10440882	0.9341719	0.11541671	0.3917919	0.85858744
GC_SKEW	0.65909415	-0.074396245	-0.004707961	-0.10784415	-0.101318955
AT_SKEW	-0.28394407	0.14454937	-0.07293732	0.2491352	-0.044612296

Table 2. Correlations between *CAI* values and *GC*-content, *GC3*, *GC*-skew and *AT*-skew values for all organisms listed in Table 1.

IV. REFERENCE SETS

Gene	Annotation	CAI	Len
MYP5240	lipoprotein VSAF (fragment)	0.84538716	1125
MYP5230	lipoprotein VSAA (fragment)	0.82216233	846
MYP4650	lipoprotein	0.7558871	1869
MYP4960	30S ribosomal protein S9	0.7448449	411
MYP5790	50S ribosomal protein L29	0.7722594	198
MYP2890	50S ribosomal protein L7/L12	0.73831713	369
MYP0760	predicted coding region	0.8143671	189

Table 3. Reference set of 7 genes for *M. pulmonis* computed by our algorithm. CAI = computed with this reference set; Len = length of the gene measured in basepairs.

Gene	Annotation	CAI	Len
TP0758	ribosomal protein S21 (rpsU)	0.7144488	210
TP0943	flagellar protein (fliS)	0.69192356	447
TP0345	phospho-N-acetylmuramoyl-pentapeptide-transferase (mraY)	0.6707904	1092
TP0490	conserved hypothetical protein	0.67239547	189
TP1033	conserved hypothetical protein	0.7363358	936
TP0847	predicted coding region	0.6556801	336
TP0777	predicted coding region	0.77040386	225
TP0420	predicted coding region	0.6868966	228
TP0248	predicted coding region	0.68390197	405
TP0942	predicted coding region	0.6738574	465

Table 4. Reference set of 10 genes for *T. pallidum* computed by our algorithm. CAI = computed with this reference set; Len = length of the gene measured in basepairs.

Gene	Annotation	CAI	Len
HP1073	copper ion binding protein (copP)	0.74117845	201
HP0337	predicted coding region	0.7436151	306
HP0756	predicted coding region	0.689386	147
HP0460	predicted coding region	0.7192915	489
HP0129	predicted coding region	0.71308863	426
HP0458	predicted coding region	0.71777135	237
HP0406	predicted coding region	0.69815063	591
HP1531	predicted coding region	0.68155533	240
HP0385	predicted coding region	0.6749625	231
HP0236	predicted coding region	0.71239513	372
HP0664	predicted coding region	0.7065184	516
HP0780	predicted coding region	0.70923877	273
HP0135	predicted coding region	0.703303	135
HP1264	predicted coding region	0.69294435	231
HP0964	predicted coding region	0.69918925	1092

Table 5. Reference set of 15 genes for *H. pylori* computed by our algorithm. CAI = computed with this reference set; gCAI = computed with the algorithm in 15 iterations; Len = length of the gene measured in basepairs. Annotation from GenBank file.

Gene	Annotation	CAI	Len
MT2972	ribosomal protein L19	0.7760756	342
MT0681	ribosomal protein L7/L12	0.76517737	393
MT1680	ribosomal protein L35	0.72539437	195
MT0713	translation elongation factor TU	0.70837396	1191
MT3568	translation initiation factor IF-1	0.6959706	222
MT0696	DNA-directed RNA polymerase, beta-prime subunit	0.6988831	3951
MT0882	thiolase	0.73571545	1212
MT3420	succinate dehydrogenase, iron-sulfur protein	0.7351174	792
MT3654	enoyl-CoA hydratase/isomerase family protein	0.71712756	744
MT2897	enoyl-CoA hydratase/isomerase family protein	0.7232695	750
MT3617	enoyl-CoA hydratase/isomerase family protein	0.7034131	792
MT0483	isocitrate lyase	0.73494107	1287
MT1697	argininosuccinate lyase	0.70521057	1413
MT2615	chorismate synthase	0.72395414	1206
MT3083	acetolactate synthase, large subunit	0.7074839	1857
MT3907	polyketide synthase	0.70167524	5202
MT1696	argininosuccinate synthase	0.7377979	1197
MT0052	1L-myo-inositol-1-phosphate synthase	0.69828033	1104
MT0397	ATP-dependent Clp protease, ATP-binding subunit ClpB	0.72395325	2547
MT2445.1	mbtH protein	0.7193337	216
MT3650	thiolase	0.71843135	1176
MT0474	aldehyde dehydrogenase	0.7150981	1524
MT3906	propionyl-CoA carboxylase, beta subunit	0.70813674	1569
MT1692	acetylglutamate kinase	0.70761764	885
MT3168	monooxygenase, flavin-binding family	0.7067472	1488
MT3346	adenosylhomocysteinase	0.70311	1488
MT3113	electron transfer flavoprotein, beta subunit	0.69501895	801
MT1691	glutamate N-acetyltransferase/ornithine acetyltransferase	0.6907989	1215
MT1322	carbonic anhydrase-related protein	0.69016165	492
MT0604	monooxygenase, putative	0.689671	1167
MT3666	substrate-CoA ligase	0.688174	1503
MT1122	cellulase-related protein	0.6878463	456
MT0592	conserved hypothetical protein	0.7367422	492
MT0550	conserved hypothetical protein	0.70107555	1590
MT2219	conserved hypothetical protein	0.7051243	867
MT0686	conserved hypothetical protein	0.7138253	156
MT2824	conserved hypothetical protein	0.7615973	753
MT1330	hypothetical protein	0.79302657	96
MT0291.3	hypothetical protein	0.7178759	195
MT0328	hypothetical protein	0.70855063	186
MT0835	hypothetical protein	0.6949745	99

Table 6. Reference set of 41 genes for *M. tuberculosis* computed by our algorithm. CAI = computed with this reference set; Len = length of the gene measured in basepairs.

Gene	Annotation	CAI	Len
hcpB	secreted protein Hcp	0.9093779	519
hcpC	secreted protein Hcp	0.8778419	519
hc	secreted protein Hcp	0.8769641	519
fpr	ferredoxin-NADP+ reductase	0.85733306	777
ty	regulatory protein TypA	0.8193153	1818
nusB	NusB protein	0.818059	480
nrdA	ribonucleoside reductase, large chain	0.8142808	2892
fbp	fructose-1,6-bisphosphatase	0.81089115	1011
ahpC	alkyl hydroperoxide reductase subunit C	0.8093085	564
sahH	S-adenosyl-L-homocysteine hydrolase	0.80415004	1410
PA5153	probable periplasmic binding protein	0.8009873	753
bfrB	bacterioferritin	0.7981742	477
guaA	GMP synthase	0.7995651	1578
PA1964	probable ATP-binding component of ABC transporter	0.79492486	1566
metK	methionine adenosyltransferase	0.79250526	1191
PA1418	probable sodium:solute symport protein	0.7919509	1392
coIII	cytochrome c oxidase, subunit III	0.7917704	888
arcB	ornithine carbamoyltransferase, catabolic	0.79073274	1011
PA3891	probable ATP-binding component of ABC transporter	0.7880887	1164
gabT	4-aminobutyrate aminotransferase	0.7857012	1281
PA1684	probable oxidase	0.7854515	546
gabD	succinate-semialdehyde dehydrogenase	0.7945166	1452
PA4022	probable aldehyde dehydrogenase	0.78778356	1521
PA1984	probable aldehyde dehydrogenase	0.78512573	1521
PA2889	probable acyl-CoA dehydrogenase	0.7843322	1161
PA0130	probable aldehyde dehydrogenase	0.7703846	1494
exaA	quinoprotein alcohol dehydrogenase	0.7699462	1872
PA5410	probable ring hydroxylating dioxygenase, alpha-subunit	0.7827083	1290
ribE	6,7-dimethyl-8-ribityllumazine synthase	0.7818556	477
glyA3	serine hydroxymethyltransferase	0.7816122	1254
ilvD	dihydroxy-acid dehydratase	0.7765565	1839
lysS	lysyl-tRNA synthetase	0.7759404	1506
leuS	leucyl-tRNA synthetase	0.76725906	2622
glnA	glutamine synthetase	0.7975705	1410
valS	valyl-tRNA synthetase	0.7950578	2853
purD	phosphoribosylamine-glycine ligase	0.77478635	1290
prpB	carboxyphosphoenolpyruvate phosphonmutase	0.77294403	897
p	inorganic pyrophosphatase	0.77080727	528
rhl	ATP-dependent RNA helicase RhlB	0.7705431	1194
speB2	"agmatinase"	0.7699218	960
PA5167	probable c4-dicarboxylate-binding protein	0.7693583	996
argG	argininosuccinate synthase	0.7693493	1218
oprH	outer membrane protein H1 precursor	0.7686601	603
comL	competence protein ComL	0.7683149	1026
PA2634	probable isocitrate lyase	0.7681374	1596
PA1574	conserved hypothetical protein	0.78250945	282
PA4673	conserved hypothetical protein	0.77488387	1101
PA5148	conserved hypothetical protein	0.78642595	273
PA2159	conserved hypothetical protein	0.79713297	417
PA0567	conserved hypothetical protein	0.8034452	159
PA4317	hypothetical protein	0.8016916	738
PA1404	hypothetical protein	0.8007509	228
PA4881	hypothetical protein	0.80839634	342
PA1830	hypothetical protein	0.7969221	315
PA3779	hypothetical protein	0.7681457	1041

Table 7. Reference set of 55 genes for *P. aeruginosa* computed by our algorithm. CAI = computed with this reference set; Len = length of the gene measured in basepairs. Annotation from GenBank file.

Gene	Annotation	CAI	Len
BB0487	ribosomal protein S17 (rpsQ)	0.8058683	255
BB0499	ribosomal protein L36 (rpmJ)	0.7937858	120
BB0266	conserved hypothetical protein	0.78943574	309
BB0848	predicted coding region	0.79416716	105
BB0138	predicted coding region	0.8685504	210
BB0150	predicted coding region	0.78167945	93
BB0609	predicted coding region	0.7624984	111
BB0433	predicted coding region	0.7345946	99

Table 8. Reference set of 8 genes for *B. burgdorferi* computed by our algorithm. CAI = computed with this reference set; Len = length of the gene measured in basepairs.

Gene	Annotation	CAI S&al.	CAI	Len
HI0544	ribosomal protein L9 (rpL9)	0.79109263	0.77064806	450
HI0641	ribosomal protein L7/L12 (rpL7/L12)	0.8026461	0.76988	372
HI1468	ribosomal protein S15 (rpS15)	0.7900408	0.7658567	270
HI0516	ribosomal protein L1 (rpL1)	0.7814362	0.7619709	690
HI0950	ribosomal protein L33 (rpL33)	0.77352995	0.74897105	171
HI1328	ribosomal protein S15 (rpS15)	0.76715857	0.7404443	270
HI1220	ribosomal protein S1 (rpS1)	0.7582906	0.73355967	1650
HI0998	ribosomal protein L34 (rpL34)	0.71783704	0.71228963	135
HI0158	ribosomal protein L32 (rpL32)	0.75685394	0.69687116	171
HI0914	elongation factor Ts (tsf)	0.7540258	0.73301697	852
HI0001	glyceraldehyde-3-phosphate dehydrogenase (gapdH)	0.7857216	0.7527914	1020
HI1370	molybdenum-pterin binding protein (mopI)	0.7631857	0.72900695	210
HI0932	enolase (eno)	0.7585627	0.7195283	1311
HI0154	acyl carrier protein (acpP)	0.7492964	0.71572727	231
HI0381	peptidoglycan-associated outer membrane lipoprotein (pal)	0.76181036	0.7385924	462
HI1164	outer membrane protein P5 (ompA)	0.73190457	0.6920485	1062
HI0017	conserved hypothetical protein	0.75034064	0.71297425	384

Table 9. Reference set of 17 genes for *H. influenzae* computed by our algorithm. The proteins listed in this table are outer-membrane proteins, and proteins involved in protein synthesis and glycolysis. CAI = computed with this reference set; CAI S&al. = computed by Perrière et Thioulouse (G. Perrière, J. Thioulouse, On-line tools for sequence retrieval and multivariate statistics in molecular biology. *Comput Appl Biosci.*, 12:63-9, 1996); Len = length of the gene measured in basepairs. Annotation from GenBank file.

Gene	Annotation	CAI	Len
rplL	50S ribosomal protein L7/L12	0.7925261	366
rplK	50S ribosomal protein L11	0.75289655	426
rplE	50S ribosomal protein L5	0.74859643	543
rplO	50S ribosomal protein L15	0.7308495	444
rplA	50S ribosomal protein L1	0.7256662	690
rpsH	30S ribosomal protein S8	0.7494694	399
rpsB	30S ribosomal protein S2	0.7815444	768
rpsD	30S ribosomal protein S4	0.81058526	612
rpsO	30S ribosomal protein S15	0.82575107	270
fusA	elongation factor G	0.7423704	2130
tuf	elongation factor Tu	0.7638007	1188
hslA	HU like DNA-binding protein	0.8057211	276
enoA	enolase (EC 4.2.1.11)	0.812126	1302
pmg	phosphoglycerate mutase (EC 5.4.2.1)	0.7793361	702
fbaA	fructose-bisphosphate aldolase (EC 4.1.2.13)	0.76547056	897
gapB	glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12)	0.85812926	1011
ptsH	phosphocarrier protein Hpr	0.7649816	267
pgk	phosphoglycerate kinase (EC 2.7.2.3)	0.755333	1197
tpiA	triosephosphate isomerase (EC 5.3.1.1)	0.7430288	759
pyk	pyruvate kinase (EC 2.7.1.40)	0.73794186	1509
acpA	acyl carrier protein	0.7291467	222
cspE	cold shock protein E	0.7446925	198

Table 10. Reference set of 22 genes for *H. lactis* computed by our algorithm. CAI = computed with this reference set; Len = length of the gene measured in basepairs. Annotation from GenBank file.

Gene	Annotation	CAI	Len
rplA	50S ribosomal protein L1	0.8333572	693
rplL	50S ribosomal protein L7/L12	0.82468474	369
rplM	50S ribosomal protein L13	0.82220304	438
rplY	50S ribosomal protein L25	0.81232023	654
rpsO	30S ribosomal protein S15	0.86341286	270
rplU	50S ribosomal protein L21	0.79048616	309
rpsD	30S ribosomal protein S4	0.8143535	603
tufA	translational elongation factor TU	0.85307634	1185
fus	translational elongation factor G	0.76435673	2082
hu	DNA-binding protein II	0.85829514	273
eno	enolase	0.83421737	1305
atpE	ATP synthase C chain	0.8315289	213
gap	glyceraldehyde-3-phosphate dehydrogenase	0.830253	1011
isaA	immunodominant antigen A	0.8291794	702
ahpC	alkyl hydroperoxide reductase subunit C	0.80512583	570
fbaA	fructose-bisphosphate aldolase	0.7880466	861
trxA	thioredoxin	0.7874468	315
tig	prolyl isomerase	0.77603126	1302
hmrB	HmrB protein	0.7665631	234
pdhD	dihydrolipoamide dehydrogenase component of pyruvate dehydrogenase E3	0.76482296	1407
ssaA	secretory antigen precursor SsaA homolog	0.76257855	804
asp23	alkaline shock protein 23	0.8680773	510
cspB	cold shock protein	0.8173901	201
SAV1739	hypothetical protein	0.7628259	492
SAV1079	hypothetical protein	0.8379392	135
SAV1845	conserved hypothetical protein	0.8194863	345
SAV0840	conserved hypothetical protein	0.8006461	195

Table 11. Reference set of 27 genes for *S. aureus* computed by our algorithm. CAI = computed with this reference set; Len = length of the gene measured in basepairs. Annotation from GenBank file.

Gene	Annotation	CAI	Len
rplL	50S ribosomal subunit protein L7/L12	0.80526525	366
rplA	50S ribosomal subunit protein L1	0.77745074	705
rpmA	50S ribosomal subunit protein L27	0.7557323	258
rplI	50s ribosomal subunit protein L9	0.7511121	450
rplY	50s ribosomal protein L25	0.7346505	285
rpmH	50s ribosomal protein l34	0.7319867	141
rplC	50S ribosomal subunit protein L3	0.7252685	630
rplO	50S ribosomal subunit protein L15	0.7225062	435
rplB	50S ribosomal subunit protein L2	0.711038	822
rpmG	50S ribosomal subunit protein L33	0.7100359	168
rpmE	50S ribosomal protein L31	0.7060832	213
rplM	50S ribosomal subunit protein L13	0.69839334	429
rpmI	50S ribosomal subunit protein L35	0.6975494	198
rplK	50S ribosomal subunit protein L11	0.69559586	429
rpsU	30S ribosomal subunit protein S21	0.72399575	216
rpsF	30s ribosomal protein S6	0.72905225	396
rpsQ	30S ribosomal subunit protein S17	0.727088	255
rpsC	30S ribosomal subunit protein S3	0.7487853	702
rpsI	30S ribosomal subunit protein S9	0.77277166	393
rpsA	30S ribosomal protein S1	0.77212965	1674
rpsB	30S ribosomal protein S2	0.75813824	726
tufA	elongation factor Tu	0.8180166	1185
tufB	elongation factor Tu	0.7812983	1185
fusA	elongation factor G	0.710757	2115
tsf	elongation factor Ts	0.7204511	852
tig	trigger factor	0.7210472	1299
hupA	histone like DNA-binding protein HU-alpha (NS2) (HU-2)	0.7028831	273
eno	enolase	0.840255	1299
gapA	glyceraldehyde 3-phosphate dehydrogenase A	0.8185587	996
fbA	fructose 1,6-bisphosphate aldolase	0.77776384	1080
groEL	GroEL protein	0.7763099	1647
pflB	formate acetyltransferase 1	0.75743234	2283
pal	peptidoglycan-associated lipoprotein precursor	0.7230564	525
glyA	serine hydroxymethyltransferase	0.7176497	1254
STY0440	probable peroxidase	0.7082626	603
sodA	manganese superoxide dismutase	0.70806855	621
acpP	acyl carrier protein	0.6969322	237
pnp	polynucleotide phosphorylase	0.6959544	2136
dnaK	DnaK protein (heat shock protein 70)	0.70241743	1917
cspA	cold shock protein	0.7871709	213
cspC	cold shock-like protein CspC	0.70469236	210
lppA	major outer membrane lipoprotein	0.830271	237
ompA	outer membrane protein A	0.744703	1053
ompC	outer membrane protein C	0.7264143	1137
STY4436	conserved hypothetical protein	0.7675389	213
yfiD	conserved hypothetical protein	0.71985114	384

Table 12. Reference set of 46 genes for *S. enterica* computed by our algorithm. CAI = computed with this reference set; Len = length of the gene measured in basepairs. Annotation from GenBank file.

Gene	Annotation	CAI S&al.	CAI	Len
rplL	ribosomal protein L12 (BL9)	0.77762175	0.77419966	372
rplA	ribosomal protein L1 (BL1)	0.7768336	0.75923043	699
rpmE	ribosomal protein L31	0.74751705	0.7409233	201
rpsP	ribosomal protein S16 (BS17)	0.6942517	0.7364216	273
rplU	ribosomal protein L21 (BL20)	0.7582789	0.7356766	309
rpmA	ribosomal protein L27 (BL24)	0.7981227	0.7132193	285
rpsB	ribosomal protein S2	0.7103496	0.7334804	741
rplM	ribosomal protein L13	0.7222421	0.73287994	438
rplT	ribosomal protein L20	0.7318787	0.7256859	360
rplB	ribosomal protein L2 (BL2)	0.70061886	0.7220637	834
rpsM	ribosomal protein S13	0.7646146	0.71814936	366
rpsI	ribosomal protein S9	0.6989476	0.71697307	393
rpsG	ribosomal protein S7 (BS7)	0.71339744	0.7130159	471
rpsQ	ribosomal protein S17 (BS16)	0.7273956	0.70621043	264
rplX	ribosomal protein L24 (BL23) (histone-like protein HPB12)	0.72415596	0.70525146	312
rpsD	ribosomal protein S4 (BS4)	0.6840494	0.70142204	603
rplP	ribosomal protein L16	0.71603644	0.68142325	435
rplN	ribosomal protein L14	0.7220372	0.6804488	369
rplJ	ribosomal protein L10 (BL5)	0.65906304	0.67901653	501
rpsF	ribosomal protein S6 (BS9)	0.6991351	0.69580835	288
fus	elongation factor G	0.6887139	0.69309187	2079
tsf	elongation factor Ts	0.7077002	0.71305716	882
tufA	elongation factor Tu	0.78300035	0.8120278	1191
abrB	transcriptional regulator	0.7581905	0.7839255	291
hbs	non-specific DNA-binding protein HBSu	0.65804386	0.69914097	279
tig	trigger factor (prolyl isomerase)	0.7057427	0.70886046	1275
ahpC	alkyl hydroperoxide reductase (small subunit)	0.76377153	0.8234251	564
sspA	small acid-soluble spore protein (alpha-type SASP)	0.85210437	0.7884815	210
gap	glyceraldehyde-3-phosphate dehydrogenase	0.7232739	0.7819248	1008
ptsH	hist-containing phosphocarrier protein of the phosphotransferase system	0.77149016	0.7349513	267
eno	enolase	0.7215345	0.7339018	1293
sspB	small acid-soluble spore protein (beta-type SASP)	0.7826275	0.70715654	204
fbaA	fructose-1,6-bisphosphate aldolase	0.7005125	0.70585805	858
trxA	thioredoxin	0.69700724	0.6991179	315
cspC	cold-shock protein	0.74345016	0.689569	201
cspD	cold-shock protein	0.7999975	0.7358131	201
cspB	major cold-shock protein	0.90190774	0.8551352	204
groEL	class I heat-shock protein (molecular chaperonin)	0.6990559	0.689498	1635
yaaK	unknown; similar to unknown proteins	0.6982412	0.7152486	324
yocJ	unknown; similar to acyl-carrier protein phosphodiesterase	0.7151327	0.7090561	627
yugI	unknown; similar to polyribonucleotide nucleotidyltransferase	0.67193806	0.6899963	393

Table 13. Reference set of 41 genes for *B. subtilis* computed by our algorithm. The proteins listed in this table are involved in folding, translation, protein synthesis, glycolysis and sporulation. CAI S&al.= computed by Sharp *et al.* (D.C. Shields, P.M. Sharp, Synonymous codon usage in *Bacillus subtilis* reflects both traditional selection and mutational biases, in *Nucleic Acids Research*, 15:8023-8040, 1987); CAI = computed with this reference set; Len = length of the gene measured in basepairs. Annotation is available at <http://pbil.univ-lyon1.fr/search/acnum.html>, at <http://genolist.pasteur.fr/SubtiList/> and from the GenBank file.

Gene	Annotation	CAI S&al.	CAI	Len
rpsU	30S ribosomal subunit protein S21	0.72601134	0.7081239	216
rpsC	30S ribosomal subunit protein S3	0.7348271	0.72319186	702
rpsI	30S ribosomal subunit protein S9	0.78217405	0.7614723	393
rpsB	30S ribosomal subunit protein S2	0.78005534	0.757977	726
rpsA	30S ribosomal subunit protein S1	0.7800922	0.7556064	1674
rplA	50S ribosomal subunit protein L1, regulates synthesis of L1 and L11	0.7728377	0.75266033	705
rplY	50S ribosomal subunit protein L25	0.762494	0.7438747	285
rpmE	50S ribosomal subunit protein L31	0.74174947	0.73376626	213
rplI	50S ribosomal subunit protein L9	0.72609437	0.71519566	450
rpmH	50S ribosomal subunit protein L34	0.73128533	0.7127593	141
rpmA	50S ribosomal subunit protein L27	0.74227947	0.7063285	258
rplC	50S ribosomal subunit protein L3	0.7132146	0.69977915	630
rplB	50S ribosomal subunit protein L2	0.71353436	0.694892	822
rpmI	50S ribosomal subunit protein A	0.71249634	0.69349897	198
rpmG	50S ribosomal subunit protein L33	0.7265544	0.69136024	168
rplK	50S ribosomal subunit protein L11	0.7072103	0.67909664	429
rplL	50S ribosomal subunit protein L7/L12	0.8448157	0.83266836	366
tsf	protein chain elongation factor EF-Ts	0.77531344	0.751327	852
tufA	protein chain elongation factor EF-Tu (duplicate of tufB)	0.8217317	0.7972499	1185
tufB	protein chain elongation factor EF-Tu (duplicate of tufA)	0.7987791	0.7706806	1185
fusA	GTP-binding protein chain elongation factor EF-G	0.7528093	0.7215798	2115
rpoC	RNA polymerase, beta prime subunit	0.70185333	0.6718387	4224
trpL	trp operon leader peptide	0.7189681	0.7070589	45
tig	trigger factor; a molecular chaperone involved in cell division	0.7383342	0.71853316	1299
eno	enolase	0.8434784	0.8224876	1299
gapA	glyceraldehyde-3-phosphate dehydrogenase A	0.839094	0.8215866	996
pflB	formate acetyltransferase 1	0.7835242	0.7610727	2283
fbA	fructose-bisphosphate aldolase, class II	0.777777	0.75712603	1080
tpiA	triosephosphate isomerase	0.74910253	0.7258386	768
pgk	phosphoglycerate kinase	0.730202	0.7084468	1164
yaiU	putative flagellin structural protein	0.71521336	0.70115334	1404
sodA	superoxide dismutase, manganese	0.72142833	0.7007647	621
ahpC	alkyl hydroperoxide reductase, C22 subunit; detoxification of hydroperoxides	0.80189043	0.78980625	564
yfiD	putative formate acetyltransferase	0.70280206	0.68787056	384
pykF	pyruvate kinase I (formerly F), fructose stimulated	0.70804596	0.684313	1413
cspA	cold shock protein 7.4, transcriptional activator of hns	0.8059303	0.7974754	213
dnaK	chaperone Hsp70; DNA biosynthesis; autoregulated heat shock proteins	0.7226461	0.69583774	1917
mopA	GroEL, chaperone Hsp60, peptide-dependent ATPase, heat shock protein	0.7961256	0.7805867	1647
ompA	outer membrane protein 3a (II*;G;d)	0.7900512	0.7678386	1041
ompC	outer membrane protein 1b (Ib;c)	0.82271796	0.8019619	1104
ompX	outer membrane protein X	0.7427497	0.72424525	516
lpp	murein lipoprotein	0.8524313	0.83522505	237

Table 14. Reference set of 42 genes for *E. coli* computed by our algorithm. The proteins listed in this table are proteins involved in translation, folding, glycolysis, protein synthesis, detoxification, and outer membrane proteins. "CAI S&al." refers to the CAI values computed by Sharp and Li (P.M. Sharp, W-H. Li., The codon adaptation index - a measure of directional synonymous codon usage bias, and its potential applications, in *Nucleic Acid Research*, 15:1281–1295, 1987); "CAI" refers to the CAI value that we compute based on this reference set; "Len" refers to the length of the gene measured in basepairs. Annotation is available at <http://genome.gen-info.osaka-u.ac.jp/bacteria/o157/search.html>, at <http://genolist.pasteur.fr/Colibri/> and from the GenBank file.

Gene	Annotation	CAI S&al.	CAI	Len
RPL39	ribosomal protein L39	0.92177784	0.91676044	156
RPL43A	ribosomal protein L37a	0.8748974	0.87354964	279
RPL30	ribosomal protein L30	0.8661489	0.8648947	318
RPL22A	ribosomal protein L22	0.8599194	0.86474013	366
RPS12	ribosomal protein S12	0.84802604	0.8520471	432
RPS6B	ribosomal protein S6	0.84681517	0.8480551	711
RPL8B	ribosomal protein L7a	0.8511839	0.8474351	771
RPL1B	60S large subunit ribosomal protein	0.8367186	0.8434909	654
RPS20	ribosomal protein S20, S22, S10	0.8297803	0.8394075	366
RPS5	ribosomal protein S5	0.84039575	0.8384518	678
RPL8A	ribosomal protein L7a	0.8439125	0.8376645	771
RPS6A	ribosomal protein S6	0.83810174	0.83664405	711
RPL5	ribosomal protein L5; for assembly of stable 60S rib. subun.	0.8314312	0.8357074	894
RPS10A	ribosomal protein S10	0.84244365	0.833947	318
RPL1A	ribosomal protein L10a, L1	0.8251024	0.8319191	654
RPL3	ribosomal protein L3	0.83519316	0.828099	1164
RPL4B	ribosomal protein L4A	0.8158778	0.82018894	1089
RPS24A	ribosomal protein S24	0.81582224	0.81839377	408
RPL32	ribosomal protein L32	0.82041085	0.81807274	393
RPL18A	ribosomal protein L18	0.81315166	0.8177621	561
RPS22A	ribosomal protein S15a	0.8158595	0.81206733	393
RPL17A	ribosomal protein L17, L22	0.81087834	0.8112943	555
RPS14A	ribosomal protein S14, S11	0.80716	0.8109136	414
RPL4A	ribosomal protein L4, L1	0.80670506	0.80942637	1089
RPP2A	ribosomal protein P2, L12eIB	0.8056242	0.8071813	321
RPS31	ribosomal protein S27a	0.81085724	0.80654216	459
RPS10B	ribosomal protein S10	0.8055274	0.8061061	318
RPS17A	ribosomal protein S17	0.81080645	0.80604094	411
RPS9B	ribosomal protein S9, S4	0.8105852	0.8056971	588
RPS2	ribosomal protein S2, S5	0.80396473	0.8024922	765
RPP0	ribosomal protein P0, L10e	0.79732054	0.79970664	939
RPL36B	ribosomal protein L36	0.8037275	0.7974957	303
RPS3	ribosomal protein S3	0.80398583	0.7974503	723
RPS23B	ribosomal protein S23, S12	0.7915378	0.7945646	438
RPL42B	ribosomal protein L36a	0.79536885	0.7911153	321
RPS19A	ribosomal protein S19	0.7980349	0.7908925	435
RPS26A	ribosomal protein S26	0.78216225	0.7891079	360
RPL15A	ribosomal protein L15	0.7875188	0.7860966	615
RPL35A	ribosomal protein L35	0.7977475	0.7858811	363
RPL9B	ribosomal protein L9	0.7808446	0.78421617	576
ASC1	40S small subunit ribosomal protein	0.7878904	0.78603286	960
RPL10	acidic ribosomal protein L10.e	0.830573	0.8335368	666
EFT2	translation elongation factor 2 (EF-2)	0.80633324	0.8065376	2529
EFT1	translation elongation factor 2 (EF-2)	0.8098462	0.8101608	2529
TEF1	translational elongation factor EF-1 alpha	0.8743042	0.87001264	1377
TEF2	translational elongation factor EF-1 alpha	0.87886	0.87437665	1377
HYP2	Translation initiation factor eIF-5A	0.8212717	0.82449794	474
TDH3	Glyceraldehyde-3-phosphate dehydrogenase 3	0.9262869	0.9272094	999
TDH2	glyceraldehyde 3-phosphate dehydrogenase	0.90739113	0.90934587	999
PDC1	pyruvate decarboxylase	0.90667886	0.9052491	1692
CDC19	Required for START A in the cell cycle and sporulation	0.8955674	0.8955937	1503
ENO1	enolase I	0.8733811	0.8719411	1314
ENO2	enolase	0.8953422	0.89357394	1314
FBA1	aldolase	0.8720781	0.87253577	1080
TDH1	Glyceraldehyde-3-phosphate dehydrogenase 1	0.85914576	0.8586988	999
TPI1	triosephosphate isomerase	0.8191473	0.82331747	747
ILV5	branched-chain amino acid biosynthesis	0.80796945	0.811605	1188
GPM1	converts 3-phosphoglycerate to 2-phosphoglycerate in glycolysis	0.81560373	0.8112557	744
ADH1	Alcohol dehydrogenase	0.81628203	0.81049323	1047
CCW12	Similarity with FL01, flocculation	0.87105066	0.8743669	402
PGK1	3-phosphoglycerate kinase	0.81739223	0.8092791	1251
SSA2	member of 70 kDa heat shock protein family	0.80456686	0.80271363	1920
SSB1	heat shock protein of HSP70 family	0.8230949	0.82478213	1842

Table 15. Reference set of 63 genes for *S. cerevisiae* computed by our algorithm. The proteins listed in this table are involved in glycolysis, translation, protein synthesis and folding. CAI S&al.= computed by Sharp & Li (P.M. Sharp, W-H. Li, The codon adaptation index - a measure of directional synonymous codon usage bias, and its potential applications, in *Nucleic Acid Research*, 15:1281–1295, 1987); CAI = computed with this reference set; Len = length of the gene measured in basepairs. Annotation is available at <http://mips.gsf.de/proj/yeast/CYGD/db/index.html> and from the GenBank file.

Gene		Annotation	CAI S&al.	CAI	Len
T28C6.6	CE03749	cuticle collagen	0.8112313	0.82438505	900
T28C6.4	CE03749	cuticle collagen	0.8112313	0.82438505	900
Y41E3.2	CE18370	cuticle collagen 1	0.77719706	0.7768235	870
W03G11.1	CE26826	cuticle collagen	0.6903886	0.7297823	840
sqt-1		cuticle collagen SQT-1	0.6920831	0.7160624	975
C34H4.4	CE27821	cuticular collagen	0.8280507	0.8246326	711
F27C1.8	CE09720	cuticular collagen	0.77852374	0.7995932	855
F30B5.1	CE04462	cuticular collagen	0.7879858	0.7912076	909
F26F12.1	CE07146	cuticular collagen	0.7676418	0.7692669	870
ZC373.7	CE02380	cuticular collagen	0.7191911	0.7475821	894
B0222.6	CE30585	collagen	0.8385449	0.8598812	852
F55C10.3	CE11182	collagen	0.7295055	0.7454781	801
T06E4.6	CE06372	collagen	0.6989508	0.7287196	873
C44C10.1	CE05408	collagen	0.7044147	0.72216415	903
F55C10.2	CE05952	collagen	0.7062916	0.7295745	900
C09G5.5	CE01485	collagen	0.73959583	0.7634808	954
F08G5.4	CE05574	collagen	0.73023075	0.75523704	900
F38A3.1	CE02213	collagen	0.71048963	0.743389	921
F23H12.4	CE05707	collagen	0.72670066	0.7436588	906
C53B4.5	CE03091	collagen	0.7282309	0.74300945	870
C29F4.1	CE03038	collagen	0.8044867	0.8171663	900
F38A3.2	CE02214	collagen	0.77697927	0.795573	921
F11G11.10	CE31783	collagen	0.7620781	0.78431	852
F57B1.3	CE11290	collagen	0.75709563	0.7820117	942
F11G11.12	CE09354	collagen	0.75496906	0.7781192	858
C09G5.4	CE01484	collagen	0.70282626	0.7205136	972
F54D1.2	CE05937	collagen	0.6766264	0.7097896	885
W05B2.1	CE20144	collagen	0.80262023	0.8212398	915
B0222.7	CE06698	collagen	0.7911819	0.81011534	885
W05B2.6	CE20148	collagen	0.7852725	0.8038523	915
col-12		collagen	0.67674595	0.7152642	951
F53F1.4	CE10938	cuticlin	0.81955147	0.8349529	318
F53F1.5	CE10940	cuticlin	0.7856933	0.8106641	648
F41F3.3	CE07217	cuticlin	0.8144529	0.80259156	471
T04C12.6	CE13148	actin	0.79776037	0.7970035	1131
act-1		actin	0.79776037	0.7970035	1131
act-4		actin	0.7980219	0.7928203	1131
T04C12.5	CE13150	actin	0.7188651	0.7277499	1131
Y75B12B.2	CE20371	Peptidyl-prolyl cis-trans isomerases	0.7764554	0.7908386	516
Y75B12B.5	CE20374	Peptidyl-prolyl cis-trans isomerases	0.7612567	0.76688236	522
C07A12.4	CE03972	protein disulfide isomerase	0.82040066	0.8367919	1482
C47E8.5	CE05441	heat shock protein (HSP90)	0.74814737	0.7749744	2109
hsp-1		heat shock 70kd protein A	0.7277199	0.7496791	1923
hsp-3		heat shock protein	0.70888317	0.72668546	1062
ost-1		osteonectin	0.737028	0.7545421	795
C47B2.3		α -tubulin	0.747564	0.7595303	1347
T22E5.5	CE04994	troponin T-like protein	0.7344623	0.75168526	1218

Table 16 (47 genes out of 164). Reference set of 164 genes for *C. elegans* computed by our algorithm. The proteins listed in this table are involved in cuticle, collagen, folding and cytoskeleton. CAI S&al. = computed by Sharp *et al.* (M. Stenico, A.T.Loyd, P.M. Sharp, Codon usage in *Caenorhabditis elegans*: delineation of translational selection and mutational biases, in *Nucleic Acid Research*, 22:2437–2446, 1994.); CAI = computed with this reference set; gCAI = computed with the algorithm in 20 iterations; Len = length of the gene measured in basepairs; Con = Confirmed or Partially Confirmed or Predicted; Exp = Experimental or Not Experimental. Annotation is available at <http://www.wormbase.org/> and from the GenBank file.

Gene		Annotation	CAI S&al.	CAI	Len
F55D10.2	CE02777	Ribosomal protein L23	0.8172343	0.83575344	444
E04A4.8	CE21392	ribosomal protein	0.86346036	0.8787172	336
F56F3.5	CE00664	Ribosomal protein S3a (human) homolog.	0.7951389	0.8115246	774
B0393.1	CE00854	40S ribosomal protein	0.7845907	0.8054699	831
D1007.12	CE09047	60S ribosomal protein L24	0.79133815	0.8101796	480
F53A3.3	CE10884	40S ribosomal protein	0.8194598	0.8261939	393
K11H12.2	CE12148	60S ribosomal protein L15	0.76940334	0.7908446	615
C23G10.3	CE01810	Ribosomal protein S3	0.7602905	0.7788229	744
F07D10.1	CE07033	ribosomal protein	0.7873541	0.78550833	591
B0513.3	CE15559	60S ribosomal protein L29	0.80664	0.81574035	189
B0041.4	CE07669	ribosomal protein L1	0.7373071	0.76077956	1038
F28C6.7a	CE03278	ribosomal protein L26 like	0.7537564	0.7722964	429
F39B2.6	CE16012	40S ribosomal protein S26	0.7511857	0.7796778	354
ZC434.2	CE06577	40S ribosomal protein S7	0.7351367	0.7582014	585
C09H10.2	CE02131	60S ribosomal protein	0.74908763	0.7662033	318
F10E7.7	CE04362	60s ribosomal protein L35A	0.73390865	0.7553734	375
T05E11.1	CE06360	40S ribosomal protein S5	0.72659427	0.75568223	633
D1007.6	CE09041	40S ribosomal protein S10	0.7295434	0.749526	450
F13B10.2	CE05598	60S ribosomal protein L3	0.74236697	0.76841605	606
F13B10.2	CE05598	60S ribosomal protein L3	0.7365042	0.73975575	600
F53G12.10	CE11024	ribosomal protein	0.72350925	0.73626524	735
F42C5.8	CE04561	40S ribosomal protein S8	0.71799725	0.7353941	627
F37C12.9	CE00821	Ribosomal protein S14	0.7016158	0.729376	459
Y57G11C.16	CE14956	ribosomal protein S13	0.7276069	0.73442966	465
Y45F10D.12	CE16650	Eukaryotic ribosomal protein L18	0.7269096	0.7366229	567
F54C9.5	CE02255	60S ribosomal protein L5	0.7086668	0.7232256	882
C32E8.2	CE08526	ribosomal protein L13	0.70585096	0.7217253	624
R11D1.8	CE06313	ribosomal protein L28 like	0.6929595	0.72906995	381
C26F1.4	CE06878	ribosomal protein/ubiquitin-like protein	0.7042706	0.72482497	393
T24B8.1	CE03709	60S ribosomal protein L32	0.7164711	0.7249689	405
B0250.1	CE18478	Ribosomal Proteins L2	0.711403	0.7159009	783
F40F11.1	CE05860	ribosomal protein S11	0.70885897	0.73149675	468
R151.3	CE00744	Ribosomal protein ML16	0.6791879	0.71294457	654
F10B5.1	CE01543	ribosomal protein L10 (QM protein)	0.7131471	0.72267926	645
F36A2.6	CE09945	40S ribosomal protein S15	0.69929785	0.7084057	456
C49H3.11	CE04237	ribosomal protein rps2	0.733767	0.7362411	819
rps-27	CE19904	Ribosomal protein S27	0.73194444	0.73764914	252
F54E7.2	CE26896	orthologous to <i>S. cerevisiae</i> rib.prot. rps12	0.7373938	0.76272	438
JC8.3		orthologous to <i>S. cerevisiae</i> rib. prot. rpl23	0.7488041	0.77381796	498
B0336.10	CE00778	orthologous to <i>S. cerevisiae</i> rib. prot. rpl23	0.77445465	0.79313916	423
Y106G6H.3	CE20413	structural protein of ribosome rpl-30	0.6733129	0.71574545	342
C06B8.8	CE20485	structural protein of ribosome rpl-38	0.7285761	0.7425733	213
C04F12.4	CE19677	structural protein of ribosome rpl-14	0.68555224	0.71065724	408
R09B3.2	CE16307	RNA recogn. motif (aka RRM/RBD/RNP dom.)	0.7326019	0.75456154	252
R09B3.3	CE16308	RNA recogn. motif (aka RRM/RBD/RNP dom.)	0.7256986	0.7537092	258
R06C1.4	CE18119	RNA recogn. motif. (aka RRM/RBD/RNP dom.)	0.78284746	0.7978317	255
R03G5.1	CE01270	elongation factor EF-1- α	0.828983	0.84455043	1392
eft-2		elong. factor Tu fam. (ATP/GTP binding P-loop)	0.6945862	0.71099484	2559
eft-3		elongation factor 1- α	0.82012135	0.83882314	1392
F54C9.1	CE02249	initiation factor 5A	0.76348066	0.776923	486
ubq-2		UBQ-2 ubiquitin; 60S rib.prot. L40	0.7182324	0.7314896	387
E04A4.7	CE16968	cytochrome C	0.7344995	0.7649009	336
F26E4.9	CE09693	cytochrome C oxidase	0.6845053	0.71674806	399
C01G6.1a	CE00863	glycerol uptake facilitator protein	0.72149706	0.74307257	873
T27E9.1	CE14263	ADP/ATP carrier protein	0.74429846	0.774402	903
unc-15		enzyme - carbohydrate metabolism	0.7441766	0.76404214	2274
unc-54		non-muscle myosin, ATP binding	0.73951143	0.73713297	2852
H28O16.1	CE18826	ATP synth. α/β subun., ATP synth. α -chain, C-term.	0.7316352	0.7499314	1617
F32D1.2	CE09866	ATP synthase ϵ -chain	0.72042197	0.7475405	165
cpr-4		cathepsin B-like cysteine proteinase	0.74153596	0.7748978	1008

Table 17 (continued, 60 genes out of 164). Reference set of 164 genes for *C. elegans* computed by our algorithm. The proteins listed in this table are involved in translation and energy. CAI S&al.= computed by Sharp *et al.*; CAI = computed with this reference set; gCAI = computed with the algorithm in 20 iterations; Len = length of the gene measured in basepairs; Con = Confirmed or Partially Confirmed or Predicted; Exp = Experimental or Not Experimental. Annotation is available at <http://www.wormbase.org/> and from the GenBank file.

Gene		Annotation	CAI S&al.	CAI	Len
gpd-2		glyceraldehyde 3-phosphate dehydrogenase	0.7699966	0.8013055	1026
gpd-3		glyceraldehyde 3-phosphate dehydrogenase	0.77775174	0.80894303	1026
F22F1.1	CE04443	histone H1	0.6802433	0.7167549	627
his-24		histone H1	0.7623889	0.7917075	627
F54E12.5	CE04501	histone H2A	0.6951529	0.7200068	384
B0035.7	CE04501	histone H2A	0.6951529	0.7200068	384
H02I12.7	CE04501	Core histones H2A, H2B, H3 and H4	0.6951529	0.7200068	384
B0035.8	CE05165	histone H2B	0.69785994	0.7235731	372
F54E12.4	CE05165	histone H2B	0.70842177	0.73359925	372
H02I12.6	CE05165	Core histones H2A, H2B, H3 and H4	0.70842177	0.73359925	372
his-62	CE05165	Core histones H2A, H2B, H3 and H4	0.7716542	0.7980976	372
ZK131.9	CE07075	histone H2B	0.6784043	0.71786237	369
ZK131.5	CE07075	histone H2B	0.6784043	0.71786237	369
F08G2.1	CE07075	Core histones H2A, H2B, H3 and H4	0.6784043	0.71786237	369
his-11		histone H2B	0.74118227	0.77650154	369
his-9		histone H3	0.6911733	0.7229097	411
Y49E10.6	CE22223	Core histone H2A/H2B/H3/H4	0.71611744	0.7240475	411
T10C6.13	CE03253	histone H3	0.703288	0.738162	411
his-17	CE03253	Core histones H2A, H2B, H3 and H4	0.70866346	0.73872226	411
his-2	CE03253	Core histone H2A/H2B/H3/H4	0.703288	0.738162	411
his-27	CE03253	Core histones H2A, H2B, H3 and H4	0.7066496	0.737655	411
F45F2.12	CE10538	histone H2B	0.6849284	0.7193837	372
his-22	CE10538	Core histones H2A, H2B, H3 and H4	0.6859228	0.722092	372
his-20	CE10538	Core histones H2A, H2B, H3 and H4	0.6859228	0.722092	372
F02A9.2	CE00133	O. volvulus 20Kd antigenic peptide	0.84111124	0.8617789	549
C06A8.3	CE02454	17k antigen (O. volvulus)	0.73442394	0.76677805	456
F25H2.11	CE09656	TCTP protein	0.7583431	0.78404456	546
F25H2.5	CE09650	nucleoside diphosphate kinase	0.78223974	0.79761213	462
F46H5.3	CE04589	arginine kinase	0.77180874	0.78561735	1134
C01G6.1b	CE19664	similarity to transport membrane protein	0.723446	0.7456615	858
F21F8.7	CE09542	protease	0.7128311	0.7446294	1170
H22K11.1	CE19495	aspartyl protease	0.7087505	0.737089	1197
K04D7.1	CE06090	guanine nucleotide-binding protein	0.71712816	0.740092	978
F52D10.3		similarity to 14-3-3 family of proteins	0.7308263	0.74877656	747
Y57G11C.15	CE14954	protein transport protein SEC61 α -subunit	0.7085232	0.72192216	1422
ZK180.5		weak simil. to CCAAT/enhancer binding prot. β	0.6953751	0.7098933	1629
C49F5.1	CE08852	s-adenosylmethionine synthetase	0.68172854	0.7106806	1212
Y38A10A.5	CE21562	calreticulin precursor	0.7025941	0.7237919	1188
F25H2.10	CE09655	deoxyribonuclease	0.7333818	0.74880517	939
vit-2		vitellogenin	0.7378615	0.76120156	4842
vit-5		structural gene for vitellogenin yp170A	0.68920606	0.7217699	4812
vit-6		vitellogenin	0.7282862	0.75441056	4956
K02F2.2	CE17154	S-adenosylhomocysteine hydrolase	0.78715014	0.8097736	1314
C27A2.2		hypothetical protein	0.7813902	0.80833983	393
F59F4.2	CE11548	unknown function	0.7734299	0.80348235	198
F09F7.2		unknown function	0.74796104	0.7556313	462
C01B10.5	CE06736	unknown function	0.7084347	0.7522071	189
C02E7.6	CE07843	unknown function	0.76233613	0.77869064	213
F44E5.1	CE18676	unknown function	0.6964583	0.7236417	237
F47B7.1	CE02743	unknown function	0.7482963	0.769293	180
F11C7.5	CE17657	unknown function	0.6918186	0.71260756	570
R102.2	CE23912	unknown function	0.73677224	0.75923604	342
B0222.8	CE06699	unknown function	0.8549517	0.881183	885
F02A9.3	CE00134	unknown function	0.8111762	0.8356803	549
C16A3.8	CE27691	unknown function	0.7735471	0.7814243	456
F46F2.3	CE05880	unknown function	0.77069974	0.77050126	414
T23F2.5	CE05000	unknown function	0.72156554	0.72576547	174

Table 18 (continued, 57 genes out of 164). Reference set of 164 genes for *C. elegans* computed by our algorithm. The proteins listed in this table are involved in energy metabolism, they are histone proteins, they cover minor roles or their function is unknown. CAI S&al.= computed by Sharp & *et al.*; CAI = computed with this reference set; Len = length of the gene measured in basepairs. Annotation is available at <http://www.wormbase.org/> and from the GenBank file.

Gene	Annotation	CAI S&al.	CAI	Len
Hsp83	Heat shock protein 83	0.7671541	0.7155633	2154
Hsc70-4	Heat shock protein cognate 4	0.75309014	0.6992676	1956
Hsp60	Heat shock protein 60	0.72596157	0.6985917	1722
Pdi	Protein disulfide isomerase	0.6974242	0.6779967	1491
CG2852	peptidyl-prolyl cis-trans isomerase	0.86208487	0.8262082	363
CG2852	peptidyl-prolyl cis-trans isomerase	0.77533513	0.7457611	618
CG5520	chaperone	0.68101126	0.6727959	2364
Cyt-c-p	Cytochrome c proximal	0.76121527	0.7262724	327
SdhB	Succinate dehydrogenase B	0.68874544	0.7176923	894
CG14028	cytochrome c oxydase respiratory chain complex IV	0.7907116	0.75225645	240
CG17280	cytochrome c oxidase respiratory chain complex IV	0.7560007	0.7549473	330
ATPsyn-d	ATP synthase, subunit d	0.6947715	0.69150233	537
CG3861	citrate (S1)-synthase mitochondrial matrix	0.678966	0.67190355	1395
Ald	Aldolase	0.8045692	0.74746174	1086
Eno	Enolase	0.80958146	0.78739625	1302
Tpi	Triose phosphate isomerase	0.7758842	0.78599674	744
Gdh	Glutamate dehydrogenase	0.72990763	0.71662116	1695
CG5548	NADH dehydrogenase	0.6685756	0.7055038	354
CG7998	malate dehydrogenase 1 mitochondrial matrix	0.73882264	0.71118003	1011
CG15093	3-hydroxyisobutyrate dehydrogenase mitochondrion	0.6853409	0.66965604	975
CG11151	estradiol 17 β -dehydrogenase	0.7258974	0.71873474	348
Adh	Alcohol dehydrogenase	0.7074499	0.67155427	771
Gapdh1	Glyceraldehyde 3 phosphate dehydrogenase 1	0.7315479	0.69419503	996
Act57B	Actin 57B	0.7989773	0.7351525	1131
Act5C	Actin 5C	0.74657106	0.695814	1131
Act79B	Actin 79B	0.71927625	0.69653267	1131
Act88F	Actin 88F	0.7137711	0.6776813	1131
α Tub84B	α -Tubulin at 84B	0.76550287	0.7088318	1353
α Tub84B	α -Tubulin at 84B	0.7603713	0.70707923	1287
β Tub56D	β -Tubulin at 56D	0.6818717	0.6727288	1344
Chd64	actin binding	0.7711419	0.7826017	528
Tm2	Tropomyosin 2	0.7154794	0.71720666	855
Mlc-c	Myosin light chain cytoplasmic	0.7185081	0.7005652	444
Prm	Paramyosin	0.6776681	0.6828633	1434
Cdle2	Cytoplasmic dynein light chain 2	0.80088216	0.789467	270
Lcp65Ac	cuticle (similarity)	0.707279	0.6727771	330
Ccp84Ad	cuticle (similarity)	0.6958904	0.6634314	600

Table 19 (37 genes out of 140). Reference set of 140 genes for *D. melanogaster* computed by our algorithm. The proteins listed in this table are involved in folding, energy, glycolysis, metabolism, cytoskeleton and cuticle. CAI S&al.= computed by Sharp *et al.* (P.M. Sharp, E. Cowe, D.G. Higgins, D.C.Shields, K.H. Wolfe, F. Wright, Codon usage patterns in *Escherichia coli*, *Bacillus subtilis*, *Saccharomices pombe*, *Drosophila melanogaster* and *Homo sapiens*; a review of the considerable within-species diversity, in *Nucleic Acids Research*, 16:8207-8211, 1988); CAI = computed with this reference set; Len = length of the gene measured in basepairs. Annotation is available at <http://flybase.bio.indiana.edu/> and from the GenBank file.

Gene	Annotation	CAI S&al.	CAI	Len
Ef1 α 48D	Elongation factor 1 α 48D	0.7678149	0.6935477	1392
Ef1 γ	involved in translational elongation	0.7474286	0.7482931	1293
CG9282	involved protein biosynthesis	0.6714627	0.6761044	468
CG9091	involved in protein biosynthesis	0.7487797	0.7276291	282
CG8332	involved in protein biosynthesis	0.72123593	0.6845774	447
CG1883	involved in protein biosynthesis	0.7497427	0.719673	585
CG3751	involved in protein biosynthesis	0.7498438	0.7170089	396
CG1475	involved in protein biosynthesis	0.74138796	0.72297066	354
CG4759	involved in protein biosynthesis	0.7177828	0.705255	408
CG4111	involved in protein biosynthesis	0.7021476	0.7025866	366
CG6846	involved in protein biosynthesis	0.76725125	0.75769067	450
RpS12	Ribosomal protein S12	0.70263416	0.6709438	420
RpS17	Ribosomal protein S17	0.729929	0.6822783	396
RpL32	Ribosomal protein L32	0.7205621	0.70275795	405
RpS18	Ribosomal protein S18	0.8171303	0.76369125	459
RpS6	Ribosomal protein S6	0.76488715	0.6951945	654
RpL40	Ribosomal protein L40	0.7415151	0.6954596	387
RpL18A	Ribosomal protein L18A	0.727202	0.7087405	534
RpL8	Ribosomal protein L8	0.7884105	0.7237671	771
RpL7A	Ribosomal protein L7A	0.7285211	0.7055514	816
RpP1	Ribosomal protein P1	0.7174073	0.67806536	342
RpL3	Ribosomal protein L3	0.7745803	0.7435334	1212
RpS9	Ribosomal protein S9	0.7346707	0.68239754	588
RpL36	Ribosomal protein L36	0.69373745	0.68364054	348
RpL13	Ribosomal protein L13	0.7129538	0.69385487	657
RpP2	Ribosomal protein P2	0.79747	0.74294674	339
RpL1	Ribosomal protein L1	0.7232364	0.6880675	1206
RpS26	Ribosomal protein S26	0.719568	0.7184378	345
RpS9	Ribosomal protein S4,	0.7346707	0.68239754	588
RpP2	60S Acidic ribosomal protein	0.79747	0.74294674	339
RpL8	Ribosomal protein L8	0.7884105	0.7237671	771
RpL36	Ribosomal protein L36	0.69373745	0.68364054	348
RpL3	Ribosomal protein L3	0.7745803	0.7435334	1212
RpL13	Ribosomal protein L13	0.7129538	0.69385487	657
CG9354	structural constituent of ribosome	0.66517556	0.6648095	507
CG11522	large ribosomal subunit (component)	0.7497474	0.7402555	789
CG11522	large ribosomal subunit (component)	0.74532866	0.73903364	732
Peritrophin-A	Peritrophin A	0.65858763	0.68103683	693
Mgstl	microsomal glutathione S-transferase-like	0.6881363	0.6706479	450
Roc1a	RING finger domain, C3HC4; smoothened receptor signaling pathway	0.69562864	0.6775361	327
CG8860	protein transporter	0.67762715	0.67232937	207
CG8579	chymotrypsin	0.697454	0.6702152	816
CG3861	citrate synthase	0.678966	0.67190355	1395
awd	abnormal wing disks	0.6889808	0.66567767	519

Table 20 (continued, 44 genes out of 140). Reference set of 140 genes for *D. melanogaster* computed by our algorithm. The proteins listed in this table are involved in translation. CAI S&al.= computed by Sharp *et al.*; CAI = computed with this reference set; Len = length of the gene measured in basepairs. Annotation is available at <http://flybase.bio.indiana.edu/> and from the GenBank file.

Gene	Annotation	CAI S&al.	CAI	Len
CG12628	glutathione transferase	0.729108	0.7106797	366
Yp1	yolk protein vitellogenesis	0.70035774	0.67395246	1320
CG14745	peptidoglycan recogn. integral plasma membrane prot.	0.6806563	0.6770313	555
Fer1HCH	Ferritin 1 heavy chain homologue	0.72571343	0.71263796	618
DebB	Developmental embryonic B	0.68211395	0.684391	255
m4	E(spl) region transcript m4	0.68610984	0.67407167	459
Vm26Aa	Vitelline membrane 26Aa	0.7383502	0.6962149	426
smt3	involved in protein-nucleus import	0.73294604	0.68222874	273
γ Try	γ Trypsin	0.71231437	0.6708819	426
blw	bellwether	0.79865503	0.73127466	1659
Sod2	Superoxide dismutase 2 (Mn)	0.656647	0.67456347	654
CG18111	odorant binding	0.8462622	0.85276777	354
CG18730	α -amylase	0.74004346	0.73913974	1482
CG14746	peptidoglycan recognition integral plasma membrane	0.7598659	0.755745	558
CG8577	peptidoglycan recognition integral plasma membrane	0.7443484	0.7355317	558
Dbi	Diazepam-binding inhibitor	0.7599149	0.77197653	261
CG8629	acyl-CoA homeostasis	0.7436277	0.72222394	255
CG8628	acyl-CoA homeostasis	0.6940497	0.67269814	255
CG15829	acyl-CoA homeostasis	0.81406534	0.7869691	249
CG5804	acyl-CoA homeostasis	0.81651604	0.77800316	249
LysC	Lysozyme C	0.76788336	0.73119533	129
GlyP	Glycogen phosphorylase	0.68173337	0.6778653	2535
Ser99Da	Serine protease 1	0.73519486	0.7089044	798
Ser99Db	Serine protease 2	0.7168793	0.6863398	759
GstD1	Glutathione S transferase D1	0.73559994	0.7285701	630
ImpL3	Ecdysone-inducible gene L3	0.6874692	0.6757043	999
oho23B	overgrown hematopoietic organs at 23B	0.80891436	0.78603673	252
Argk	Arginine kinase	0.8114714	0.7674014	1071
Lsp1 β	Larval serum protein 1 β	0.83539486	0.83400327	2370
Mp20	Muscle protein	0.79795563	0.7595253	555
Amy-d	Amylase distal	0.7375889	0.73895395	1482
Ser4	Serine protease 4	0.7849072	0.73946166	801
Rack1	Receptor of activated protein kinase C 1	0.74969864	0.72999686	957
CG8869	chymotrypsin	0.7654211	0.7349438	831
CG8871	chymotrypsin	0.7633865	0.70306724	777
CG7118	serine-type endopeptidase	0.74009347	0.6956582	783
CG6467	serine-type endopeptidase	0.6935264	0.6684057	816
CG5177	trehalase phosphatase	0.70744395	0.68405604	831
PebIII	Ejaculatory bulb protein III	0.69698715	0.7070293	375
Jafrac1	thioredoxin peroxidase 1	0.70550215	0.6916231	585
CG3731	mitochondrial processing peptidase (complex)	0.6856974	0.6747257	1149
CG6295	triacylglycerol lipase	0.70779693	0.66878366	1017
CG6783	unknown function	0.82480294	0.7981521	393
CG18594	unknown function	0.80518544	0.7779789	531
CG4800	unknown function	0.78204465	0.7535077	519
CG6770	unknown function	0.7826638	0.7673293	210
CG7710	unknown function	0.7315584	0.74209875	285
CG16885	unknown function	0.7681256	0.71733844	780
CG11086	unknown function	0.7193182	0.7136454	492
CG13551	unknown function	0.72908616	0.74630606	138
CG12726	unknown function	0.7104443	0.72009337	621
CG5468	unknown function	0.7420228	0.6718044	867
CG15704	unknown function	0.7149938	0.73556787	156
CG8756	unknown function	0.6685625	0.6725142	1596
CG1153	unknown function	0.6818915	0.66592556	867
CG2297	unknown function	0.7412106	0.7446065	432
CG7592	unknown function	0.7420241	0.72597927	450
CG14770	unknown function	0.66770256	0.6858921	456
CG14235	unknown function	0.6459286	0.6862359	279

Table 21 (continued, 59 genes out of 140). Reference set of 140 genes for *D. melanogaster* computed by our algorithm. CAI S&al.= computed by Sharp *et al.* ; CAI = computed with this reference set; Len = length of the gene measured in basepairs. Annotation is available at <http://flybase.bio.indiana.edu/> and from the GenBank file.