

I. ACCESSION NUMBERS

Organism	Accession Number in GeneBank FTP
<i>Mycoplasma pulmonis</i>	AL445566
<i>Mycobacterium tuberculosis</i>	AE000516
<i>Treponema pallidum</i>	AE000520
<i>Helicobacter pylori</i>	AE000511
<i>Haemophilus influenzae</i> Rd	L42023
<i>Salmonella enterica</i>	AL513382
<i>Staphylococcus aureus</i>	BA00017
<i>Escherichia coli</i>	U00096
<i>Bacillus subtilis</i>	AL009126
<i>Lactococcus lactis</i>	AE005176
<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. FUNCTIONAL CATEGORIES

For each organism, each functional category is given with its corresponding number of ORFs. Notice that the same ORF, when it corresponds to a protein with several functions, is counted several times.

Functional Categories of *H. pylori*

- 01 Amino acid biosynthesis (41)
- 02 Purines, pyrimidines, nucleosides, and nucleotides (35)
- 03 Fatty acid and phospholipid metabolism (29)
- 04 Biosynthesis of cofactors, prosthetic groups, and carriers (57)
- 05 Central intermediary metabolism, energy metabolism (125)
- 06 Transport and binding proteins (88)
- 07 DNA metabolism (86)
- 08 Transcription (15)
- 09 Protein synthesis (99)
- 10 Protein fate (46)
- 11 Regulatory functions (21)
- 12 Cell envelope (108)
- 13 Cellular processes (115)
- 14 Other categories (16)
- 15 Unknown and hypothetical (695)

Functional Categories of *B. subtilis*

- 01 Metabolism of Amino-acids and Related Molecules (169)
- 02 Metabolism of Nucleotides and Nucleic Acids (71)
- 03 Metabolism of Lipids (68)

- 04 Metabolism of Co-enzymes and Prosthetic Groups (86)
- 05 Metabolism of Phosphate, Metabolism of Sulphur, Membrane Bioenergetics (Electron Transport Chain and ATP Synthetase), Metabolism of Carbohydrates and Related Molecules: Specific Pathways, Main Glycolytic Pathways, TCA Cycle (333)
- 06 Transport/Binding Proteins and Lipoproteins (363)
- 07 DNA Replication, Restriction/Modification and Repair, Recombination, Packaging and Segregation; Phage-related Functions; Transposon and IS (177)
- 08 RNA Synthesis: Initiation, Elongation, Termination, Modification (49)
- 09 Protein Synthesis: Ribosomal Proteins, Aminoacyl-tRNA Synthetase, Initiation, Elongation, Termination (96)
- 10 Protein Modification, Protein Folding, Protein Secretion (51)
- 11 RNA regulation, Adaptation to Atypical Conditions (275)
- 12 Cell Envelope and Cellular Processes: Cell Wall, Sensors (Signal Transduction) (140)
- 13 Motility and Chemotaxis, Cell division, Sporulation, Germination, Transformation/Competence, Detoxification (314)
- 14 Antibiotic Production, Miscellaneous (53)
- 15 Similar to Unknown Proteins, No similarity (1534)

Functional Categories of *E. coli*

- 01 Amino-acids Biosynthesis (68)
- 02 Nucleotide Biosynthesis (22)
- 03 Fatty Acids Biosynthesis (16)
- 04 Biosynthesis of Co-factors (67)
- 05 Central Intermediary Metabolism, Degradation of Small Molecules, Energy Metabolism, Carbon (290)
- 06 Transport/Binding Proteins (239)
- 07 DNA Synthesis, Modification, Degradation; Laterally Acquired Elements (101)
- 08 RNA Synthesis, Modification, Degradation (18)
- 09 Protein Synthesis, Modification, Degradation, Ribosome Constituent (59)
- 10 Folding and Ushering Proteins (49)
- 11 Global Regulatory Functions, Adaptation (78)
- 12 Cell Envelop (66)
- 13 Cell Division, Protection Responses (57)
- 14 Not Classified (393)
- 15 Unknown function (602)

Functional Categories of *S. cerevisiae*

- 01 Metabolism (1066)
- 02 Energy (252)
- 03 Cellular Transport and Transport Mechanisms (495)
- 04 Transport Facilitation (313)
- 05 Cell Cycle and DNA processing (628)
- 06 Transposable Elements, Viral and Plasmid Proteins (116)
- 07 Transcription (771)
- 08 Protein Synthesis (359)
- 09 Protein Fate (folding, modification, destination) (595)
- 10 Protein Activity Regulation (13)
- 11 Protein with Binding Function or Cofactor Requirements (4)
- 12 Cell Fate (427)
- 13 Control of Cellular Organisation (209)
- 14 Cellular Communication/Signal Transduction Mechanisms (59)
- 15 Cell Rescue, Defense and Virulence (278)
- 16 Regulation of/Interaction with Cellular Environment (199)
- 17 Classification not yet clear-cut (115)
- 18 Unclassified Proteins (2399)

Localization Categories of *S. cerevisiae*

- S01 Cell Wall (38)
- S02 Centrosome (31)
- S03 Chromosome (44)
- S04 Cytoplasm (554)
- S05 Cytoskeleton (108)
- S06 Endoplasmic reticulum (157)
- S07 Endosome (12)
- S08 Extracellular secretion proteins (20)
- S09 Golgi (82)
- S10 Intracellular transport vesicles (42)
- S11 Mitochondrion (366)
- S12 Nucleus (774)
- S13 Peroxisome (39)
- S14 Plasma membrane (145)
- S15 Prokaryotic cell membrane (1)
- S16 Vacuole (59)
- S17 Other subcellular localisations (8)

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

ttt	1.000	ttc	0.099	ttg	0.585	tta	1.000	tct	0.796	tcc	0.082	tcg	0.041	tca	0.163
tgt	0.083	tgc	1.000	tgg	1.000	tga	0.000	tat	1.000	tac	0.379	tag	0.000	taa	0.000
ctt	0.287	ctc	0.117	ctg	0.011	cta	0.043	cct	1.000	ccc	0.125	ccg	0.083	cca	0.250
cgt	0.435	cgc	0.826	cgg	0.000	cga	0.217	cat	1.000	cac	0.091	cag	0.022	caa	1.000
ggt	0.511	gtc	0.289	gtg	1.000	gta	0.111	gct	1.000	gcc	0.250	gcg	0.904	gca	0.192
ggg	0.333	ggc	0.852	ggg	1.000	gga	0.111	gat	1.000	gac	0.213	gag	0.283	gaa	1.000
att	1.000	atc	0.766	atg	1.000	ata	0.203	act	0.650	acc	1.000	acg	0.500	aca	0.400
agt	0.265	agc	1.000	agg	0.783	aga	1.000	aat	1.000	aac	0.758	aag	0.242	aaa	1.000

Weights for *H. pylori*.

ttt	0.515	ttc	1.000	ttg	0.049	tta	1.000	tct	1.000	tcc	0.000	tcg	0.000	tca	0.350
tgt	1.000	tgc	0.250	tgg	1.000	tga	0.000	tat	0.421	tac	1.000	tag	0.000	taa	0.000
ctt	0.191	ctc	0.000	ctg	0.005	cta	0.034	cct	0.322	ccc	0.000	ccg	0.000	cca	1.000
cgt	1.000	cgc	0.122	cgg	0.000	cga	0.000	cat	0.200	cac	1.000	cag	0.019	caa	1.000
ggt	1.000	gtc	0.029	gtg	0.188	gta	0.747	gct	0.538	gcc	0.017	gcg	0.210	gca	1.000
ggg	1.000	ggc	0.302	ggg	0.005	gga	0.023	gat	1.000	gac	0.398	gag	0.084	gaa	1.000
att	0.512	atc	1.000	atg	1.000	ata	0.000	act	1.000	acc	0.099	acg	0.076	aca	0.176
agt	0.136	agc	0.184	agg	0.000	aga	0.014	aat	0.480	aac	1.000	aag	0.027	aaa	1.000

Weights for *H. influenzae*.

ttt	0.370	ttc	1.000	ttg	0.566	tta	0.030	tct	0.423	tcc	0.000	tcg	0.000	tca	1.000
tgt	1.000	tgc	0.200	tgg	1.000	tga	0.000	tat	0.218	tac	1.000	tag	0.000	taa	0.000
ctt	1.000	ctc	0.166	ctg	0.000	cta	0.000	cct	0.423	ccc	0.000	ccg	0.007	cca	1.000
cgt	1.000	cgc	0.171	cgg	0.000	cga	0.000	cat	0.368	cac	1.000	cag	0.000	caa	1.000
ggt	1.000	gtc	0.081	gtg	0.036	gta	0.363	gct	1.000	gcc	0.044	gcg	0.065	gca	0.530
ggg	1.000	ggc	0.071	ggg	0.008	gga	0.269	gat	1.000	gac	0.940	gag	0.006	gaa	1.000
att	0.340	atc	1.000	atg	1.000	ata	0.000	act	1.000	acc	0.009	acg	0.000	aca	0.594
agt	0.023	agc	0.080	agg	0.000	aga	0.005	aat	0.292	aac	1.000	aag	0.027	aaa	1.000

Weights for *L. lactis*.

ttt	0.414	ttc	1.000	ttg	0.009	tta	1.000	tct	0.407	tcc	0.000	tcg	0.000	tca	1.000
tgt	1.000	tgc	0.333	tgg	1.000	tga	0.000	tat	0.700	tac	1.000	tag	0.000	taa	0.000
ctt	0.108	ctc	0.000	ctg	0.000	cta	0.026	cct	0.483	ccc	0.000	ccg	0.017	cca	1.000
cgt	1.000	cgc	0.081	cgg	0.000	cga	0.011	cat	0.667	cac	1.000	cag	0.012	caa	1.000
gtt	1.000	gtc	0.018	gtg	0.040	gta	0.806	gct	1.000	gcc	0.011	gcg	0.114	gca	0.849
ggg	1.000	ggc	0.113	ggg	0.003	gga	0.176	gat	1.000	gac	0.800	gag	0.068	gaa	1.000
att	0.848	atc	1.000	atg	1.000	ata	0.010	act	1.000	acc	0.005	acg	0.023	aca	0.690
agt	0.167	agc	0.185	agg	0.000	aga	0.156	aat	0.667	aac	1.000	aag	0.026	aaa	1.000

Weights for *S. aureus*.

ttt	0.275	ttc	1.000	ttg	0.018	tta	0.016	tct	1.000	tcc	0.819	tcg	0.029	tca	0.021
tgt	0.288	tgc	1.000	tgg	1.000	tga	0.000	tat	0.298	tac	1.000	tag	0.000	taa	0.000
ctt	0.029	ctc	0.030	ctg	1.000	cta	0.004	cct	0.089	ccc	0.003	ccg	1.000	cca	0.104
cgt	1.000	cgc	0.469	cgg	0.000	cga	0.002	cat	0.211	cac	1.000	cag	1.000	caa	0.109
gtt	1.000	gtc	0.167	gtg	0.395	gta	0.443	gct	1.000	gcc	0.284	gcg	0.721	gca	0.444
ggg	0.897	ggc	1.000	ggg	0.022	gga	0.008	gat	0.495	gac	1.000	gag	0.305	gaa	1.000
att	0.268	atc	1.000	atg	1.000	ata	0.002	act	0.577	acc	1.000	acg	0.108	aca	0.046
agt	0.033	agc	0.469	agg	0.000	aga	0.000	aat	0.098	aac	1.000	aag	0.228	aaa	1.000

Weights for *S. enterica*.

ttt	0.317	ttc	1.000	ttg	0.138	tta	0.387	tct	1.000	tcc	0.081	tcg	0.009	tca	0.314
tgt	0.786	tgc	1.000	tgg	1.000	tga	0.000	tat	0.292	tac	1.000	tag	0.000	taa	0.000
ctt	1.000	ctc	0.040	ctg	0.099	cta	0.099	cct	0.885	ccc	0.015	ccg	0.267	cca	1.000
cgt	1.000	cgc	0.544	cgg	0.000	cga	0.000	cat	0.600	cac	1.000	cag	0.187	caa	1.000
gtt	1.000	gtc	0.130	gtg	0.176	gta	0.611	gct	1.000	gcc	0.057	gcg	0.291	gca	0.551
ggg	1.000	ggc	0.774	ggg	0.030	gga	0.660	gat	1.000	gac	0.905	gag	0.219	gaa	1.000
att	0.383	atc	1.000	atg	1.000	ata	0.006	act	1.000	acc	0.005	acg	0.167	aca	0.964
agt	0.045	agc	0.260	agg	0.000	aga	0.075	aat	0.168	aac	1.000	aag	0.110	aaa	1.000

Weights for *B. subtilis*.

ttt	0.199	ttc	1.000	ttg	0.016	tta	0.014	tct	1.000	tcc	0.716	tcg	0.028	tca	0.032
tgt	0.373	tgc	1.000	tgg	1.000	tga	0.000	tat	0.238	tac	1.000	tag	0.000	taa	0.000
ctt	0.024	ctc	0.035	ctg	1.000	cta	0.002	cct	0.085	ccc	0.003	ccg	1.000	cca	0.141
cgt	1.000	cgc	0.409	cgg	0.000	cga	0.004	cat	0.237	cac	1.000	cag	1.000	caa	0.107
gtt	1.000	gtc	0.078	gtg	0.246	gta	0.508	gct	1.000	gcc	0.144	gcg	0.462	gca	0.550
ggg	1.000	ggc	0.643	ggg	0.015	gga	0.004	gat	0.423	gac	1.000	gag	0.200	gaa	1.000
att	0.209	atc	1.000	atg	1.000	ata	0.002	act	0.750	acc	1.000	acg	0.045	aca	0.036
agt	0.025	agc	0.489	agg	0.000	aga	0.000	aat	0.051	aac	1.000	aag	0.213	aaa	1.000

Weights for *E. coli*.

ttt	0.148	ttc	1.000	ttg	1.000	tta	0.127	tct	1.000	tcc	0.700	tcg	0.000	tca	0.013
tgt	1.000	tgc	0.060	tgg	1.000	tga	0.000	tat	0.079	tac	1.000	tag	0.000	taa	0.000
ctt	0.005	ctc	0.000	ctg	0.002	cta	0.039	cct	0.083	ccc	0.003	ccg	0.000	cca	1.000
cgt	0.147	cgc	0.000	cgg	0.000	cga	0.000	cat	0.220	cac	1.000	cag	0.003	caa	1.000
gtt	1.000	gtc	0.813	gtg	0.018	gta	0.001	gct	1.000	gcc	0.302	gcg	0.002	gca	0.006
ggg	1.000	ggc	0.018	ggg	0.002	gga	0.003	gat	0.622	gac	1.000	gag	0.011	gaa	1.000
att	0.813	atc	1.000	atg	1.000	ata	0.002	act	0.992	acc	1.000	acg	0.000	aca	0.017
agt	0.026	agc	0.031	agg	0.000	aga	1.000	aat	0.057	aac	1.000	aag	1.000	aaa	0.092

Weights for *S. cerevisiae*.

ttt	0.060	ttc	1.000	ttg	0.296	tta	0.005	tct	0.647	tcc	1.000	tcg	0.148	tca	0.199
tgt	0.198	tgc	1.000	tgg	1.000	tga	0.000	tat	0.156	tac	1.000	tag	0.000	taa	0.000
ctt	0.860	ctc	1.000	ctg	0.048	cta	0.004	cct	0.012	ccc	0.006	ccg	0.012	cca	1.000
cgt	1.000	cgc	0.636	cgg	0.006	cga	0.039	cat	0.375	cac	1.000	cag	0.363	caa	1.000
gtt	0.678	gtc	1.000	gtg	0.134	gta	0.047	gct	0.849	gcc	1.000	gcg	0.015	gca	0.130
ggg	0.057	ggc	0.021	ggg	0.006	gga	1.000	gat	0.838	gac	1.000	gag	1.000	gaa	0.417
att	0.293	atc	1.000	atg	1.000	ata	0.005	act	0.418	acc	1.000	acg	0.024	aca	0.067
agt	0.037	agc	0.192	agg	0.010	aga	0.497	aat	0.190	aac	1.000	aag	1.000	aaa	0.090

Weights for *C. elegans*.

ttt	0.055	ttc	1.000	ttg	0.138	tta	0.002	tct	0.135	tcc	1.000	tcg	0.408	tca	0.033
tgt	0.092	tgc	1.000	tgg	1.000	tga	0.000	tat	0.124	tac	1.000	tag	0.000	taa	0.000
ctt	0.057	ctc	0.234	ctg	1.000	cta	0.018	cct	0.107	ccc	1.000	ccg	0.149	cca	0.155
cgt	0.558	cgc	1.000	cgg	0.032	cga	0.027	cat	0.194	cac	1.000	cag	1.000	caa	0.078
gtt	0.232	gtc	0.709	gtg	1.000	gta	0.030	gct	0.247	gcc	1.000	gcg	0.065	gca	0.036
ggg	0.551	ggc	1.000	ggg	0.005	gga	0.365	gat	0.611	gac	1.000	gag	1.000	gaa	0.089
att	0.238	atc	1.000	atg	1.000	ata	0.008	act	0.125	acc	1.000	acg	0.105	aca	0.036
agt	0.042	agc	0.456	agg	0.084	aga	0.016	aat	0.125	aac	1.000	aag	1.000	aaa	0.047

Weights for *D. melanogaster*.

IV. REFERENCE SETS

Gene	Annotation	CAI	gCAI	Len
HP0337	predicted coding region	0.7436151	0.62494326	306
HP0756	predicted coding region	0.689386	0.60070705	147
HP1073	copper ion binding protein (copP)	0.74117845	0.58820647	201
HP0460	predicted coding region	0.7192915	0.5833216	489
HP0129	predicted coding region	0.71308863	0.58144265	426
HP0458	predicted coding region	0.71777135	0.57769805	237
HP0406	predicted coding region	0.69815063	0.57406867	591
HP1531	predicted coding region	0.68155533	0.5740182	240
HP0385	predicted coding region	0.6749625	0.56579065	231
HP0236	predicted coding region	0.71239513	0.5635466	372
HP0664	predicted coding region	0.7065184	0.55820733	516
HP0780	predicted coding region	0.70923877	0.5518484	273
HP0135	predicted coding region	0.703303	0.5489145	135
HP1264	predicted coding region	0.69294435	0.54316443	231
HP0964	predicted coding region	0.69918925	0.54185027	1092

Table 2. Reference set of 17 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	gCAI	Len
HI0641	ribosomal protein L7/L12 (rpL7/L12)	0.74424595	0.7061593	372
HI0544	ribosomal protein L9 (rpL9)	0.7504977	0.69901454	450
HI0516	ribosomal protein L1 (rpL1)	0.74634653	0.6794397	690
HI1468	ribosomal protein S15 (rpS15)	0.7376159	0.67339945	270
HI1220	ribosomal protein S1 (rpS1)	0.7289901	0.6476252	1650
HI1328	ribosomal protein S15 (rpS15)	0.71337193	0.64662886	270
HI0950	ribosomal protein L33 (rpL33)	0.68933946	0.6234312	171
HI1443	ribosomal protein L13 (rpL13)	0.67388964	0.5906101	429
HI0914	elongation factor Ts (tsf)	0.7227573	0.6508328	852
HI0001	glyceraldehyde-3-phosphate dehydrogenase (gapdH)	0.7434514	0.6612865	1020
HI0932	enolase (eno)	0.71558255	0.6377796	1311
HI1370	molybdenum-pterin binding protein (mopI)	0.6913084	0.6269895	210
HI0154	acyl carrier protein (acpP)	0.6787583	0.62124103	231
HI0381	peptidoglycan-associated outer membrane lipoprotein (pal)	0.7166589	0.6430323	462
HI1164	outer membrane protein P5 (ompA)	0.6962986	0.6032673	1062
HI0017	conserved hypothetical protein	0.69823754	0.5933235	384

Table 3. Reference set of 16 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; gCAI = computed with the algorithm in 15 iterations; Len = length of the gene measured in basepairs. Annotation from GenBank file.

Gene	Annotation	CAI	gCAI	Len
rpsD	30S ribosomal protein S4	0.79232126	0.75940543	612
rpsO	30S ribosomal protein S15	0.7829348	0.75106317	270
rpsB	30S ribosomal protein S2	0.7673643	0.7275176	768
rpsH	30S ribosomal protein S8	0.7245862	0.68962955	399
rpsG	30S ribosomal protein S7	0.713859	0.66294825	468
rplL	50S ribosomal protein L7/L12	0.7614957	0.7468566	366
rplE	50S ribosomal protein L5	0.7296806	0.6881821	543
rplK	50S ribosomal protein L11	0.7255184	0.6893408	426
rplA	50S ribosomal protein L1	0.7103788	0.66986305	690
rplO	50S ribosomal protein L15	0.70916504	0.6739384	444
rplQ	50S ribosomal protein L17	0.69242036	0.6558399	381
tuf	elongation factor Tu	0.754167	0.7012061	1188
fusA	elongation factor G	0.7383873	0.68764657	2130
hslA	HU like DNA-binding protein	0.76503164	0.74229693	276
pmg	phosphoglycerate mutase	0.7644217	0.7193454	702
gapB	glyceraldehyde 3-phosphate dehydrogenase	0.8453794	0.8133781	1011
enoA	enolase	0.8019261	0.7655435	1302
fbaA	fructose-bisphosphate aldolase	0.75477535	0.6927456	897
pyk	pyruvate kinase	0.73128366	0.6882763	1509
pgk	phosphoglycerate kinase	0.7452723	0.70852005	1197
tpiA	triosephosphate isomerase	0.72852665	0.66270983	759
ptsH	phosphocarrier protein Hpr	0.72346354	0.69375473	267
ldh	L-lactate dehydrogenase	0.7175333	0.6621635	978

Table 4. Reference set of 23 genes for *H. lactis* 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	gCAI	Len
rpsO	30S ribosomal protein S15	0.81610066	0.72877306	270
rpsD	30S ribosomal protein S4	0.7959033	0.70073026	603
rpsB	30S ribosomal protein S2	0.756387	0.66112363	768
rpsL	30S ribosomal protein S9	0.73311776	0.6360313	393
rplY	50S ribosomal protein L25	0.7929298	0.7239046	654
rplL	50S ribosomal protein L7/L12	0.79209507	0.7406857	369
rplM	50S ribosomal protein L13	0.7864378	0.69562966	438
rplA	50S ribosomal protein L1	0.8125852	0.73525614	693
rplU	50S ribosomal protein L21	0.7499876	0.67493534	309
rplT	50S ribosomal protein L20	0.735607	0.62590337	357
tufA	translational elongation factor TU	0.8427333	0.7571113	1185
fus	translational elongation factor G	0.76689804	0.6611207	2082
SAV1257	elongation factor TS	0.74721575	0.65969044	882
hu	DNA-binding protein II	0.8100686	0.7558042	273
asp23	alkaline shock protein 23	0.8375831	0.7765389	510
cspB	cold shock protein	0.7486226	0.6742806	201
eno	enolase	0.82303995	0.7370553	1305
gap	glyceraldehyde-3-phosphate dehydrogenase	0.8177097	0.7341422	1011
isaA	immunodominant antigen A	0.7970232	0.6928569	702
ahpC	alkyl hydroperoxide reductase subunit C	0.78206515	0.6690921	570
fbaA	fructose-bisphosphate aldolase	0.7717179	0.6706211	861
atpE	ATP synthase C chain	0.76903063	0.71816915	213
tig	prolyl isomerase	0.76505184	0.68298984	1302
pdhD	dihydrolipoamide dehydrogenase component of pyruvate dehydrogenase E3	0.7513818	0.6579734	1407
trxA	thioredoxin	0.74493504	0.6257628	315
pflB	formate acetyltransferase	0.7427024	0.6302557	2250
SAV1845	conserved hypothetical protein	0.7846938	0.71940756	345
SAV1079	hypothetical protein	0.7479107	0.69792646	135

Table 5. Reference set of 28 genes for *S. aureus* 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	gCAI	Len
STY0981	30S ribosomal protein S1	0.77020264	0.71795464	1674
STY3524	30S ribosomal subunit protein S9	0.7513097	0.70432645	393
STY0239	30S ribosomal protein S2	0.74677426	0.6895201	726
rpsC	30S ribosomal subunit protein S3	0.73768336	0.6838237	702
rpsF	30s ribosomal protein S6	0.71039766	0.65546745	396
rpsQ	30S ribosomal subunit protein S17	0.69768554	0.64266616	255
STY3388	30S ribosomal subunit protein S21	0.68831396	0.64747363	216
STY3733	50S ribosomal subunit protein L7/L12	0.7807467	0.7482504	366
STY3735	50S ribosomal subunit protein L1	0.7687311	0.7257433	705
STY2461	50s ribosomal protein L25	0.7054185	0.6602064	285
rplO	50S ribosomal subunit protein L15	0.7036093	0.6526472	435
rplB	50S ribosomal subunit protein L2	0.7013877	0.64084065	822
STY3736	50S ribosomal subunit protein L11	0.6840696	0.62761045	429
rplD	50S ribosomal subunit protein L4	0.68035567	0.62079966	606
STY3525	50S ribosomal subunit protein L13	0.67814904	0.6202209	429
rplI	50s ribosomal subunit protein L9	0.7325627	0.6912163	450
STY3482	50S ribosomal subunit protein L27	0.7176312	0.6611548	258
rplC	50S ribosomal subunit protein L3	0.71220684	0.65950775	630
tufA	elongation factor Tu	0.81097806	0.76226616	1185
tufB	elongation factor Tu	0.7755275	0.72154164	1185
STY0240	elongation factor Ts	0.71670246	0.6737086	852
fusA	elongation factor G	0.7105377	0.6395741	2115
cspA	cold shock protein	0.7408079	0.6928547	213
STY0012	DnaK protein (heat shock protein 70)	0.70422256	0.64647484	1917
STY1745	major outer membrane lipoprotein	0.7868232	0.748067	237
STY1091	outer membrane protein A	0.738334	0.6725191	1053
ompC	outer membrane protein C	0.7235134	0.6542158	1137
STY3081	enolase	0.8375476	0.7932138	1299
STY1825	glyceraldehyde 3-phosphate dehydrogenase A	0.8114941	0.7633586	996
groEL	GroEL protein	0.7776223	0.7354828	1647
STY3226	fructose 1,6-bisphosphate aldolase	0.77299035	0.7147705	1080
STY0973	formate acetyltransferase 1	0.7577766	0.6961704	2283
STY0489	trigger factor	0.7213502	0.6689904	1299
STY2802	serine hydroxymethyltransferase	0.72041327	0.66239303	1254
STY0795	peptidoglycan-associated lipoprotein precursor	0.71348023	0.65124506	525
STY3816	manganese superoxide dismutase	0.7016906	0.631047	621
STY0440	probable peroxidase	0.7003627	0.64333695	603
STY3463	polynucleotide phosphorylase	0.69970924	0.6387803	2136
STY3648	ketol-acid reductoisomerase	0.69317985	0.61908185	1476
STY3227	phosphoglycerate kinase	0.6908919	0.63209015	1164
STY0175	pyruvate dehydrogenase E1 component	0.68690586	0.6081711	2664
STY0782	succinyl-CoA synthetase α -chain	0.68670934	0.615398	870
ppa	inorganic pyrophosphatase	0.68375033	0.6155168	531
STY0177	dihydrolipoamide dehydrogenase	0.6764436	0.60226446	1428
STY4436	conserved hypothetical protein	0.7219101	0.6628953	213
STY2839	conserved hypothetical protein	0.70165807	0.63994414	384

Table 6. Reference set of 46 genes for *S. enterica* 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 S&al.	CAI	gCAI	Len
cspB	major cold-shock protein	0.84329706	0.80474055	0.7664587	204
cspD	cold-shock protein	0.74860716	0.68984735	0.6264418	201
groEL	class I heat-shock protein (chaperonin)	0.6936195	0.69467604	0.6239956	1635
dnaK	class I heat-shock protein (molecular chaperone)	0.65921754	0.6703921	0.5841617	1836
tufA	elongation factor Tu	0.7744256	0.8096196	0.7385357	1191
tsf	elongation factor Ts	0.6974866	0.712357	0.63843256	882
abrB	transcriptional pleiotropic regulator of transition state genes	0.7247652	0.7536265	0.68375885	291
hbs	non-specific DNA-binding protein HBSu	0.62876916	0.67245173	0.6057464	279
tig	trigger factor (prolyl isomerase)	0.698693	0.70866734	0.64291817	1275
ahpC	alkyl hydroperoxide reductase (small subunit)	0.74626684	0.80985093	0.727566	564
sspB	small acid-soluble spore protein (major β -type SASP)	0.7333193	0.6682616	0.6029736	204
sspA	small acid-soluble spore protein (major α -type SASP)	0.7989409	0.7440473	0.7089121	210
trxA	thioredoxin	0.66913515	0.68329763	0.5905421	315
eno	enolase	0.71439034	0.7361227	0.6568578	1293
fbaA	fructose-1,6-bisphosphate aldolase	0.6901457	0.70430297	0.62257797	858
ptsH	histidine-containing phosphocarrier protein of the PTS (HPr prot.)	0.73431677	0.70632434	0.64166486	267
gap	glyceraldehyde-3-phosphate dehydrogenase	0.71408945	0.7782096	0.7094212	1008
fus	Elongation factor G	0.68451464	0.6963889	0.60570395	2079
yaaK	unknown; similar to unknown proteins	0.67107606	0.6898711	0.63752186	324
yocJ	unknown; similar to acyl-carrier protein phosphodiesterase	0.70060194	0.703833	0.6247659	627
yugI	unknown; similar to polyribonucleotide nucleotidyltransferase	0.65053827	0.67704433	0.6059066	393
rplA	ribosomal protein L1 (BL1)	0.76239496	0.75131774	0.6662966	699
rpsB	ribosomal protein S2	0.69814533	0.72752005	0.6600376	741
rpsP	ribosomal protein S16 (BS17)	0.66230136	0.7106838	0.649713	273
rplU	ribosomal protein L21 (BL20)	0.7267738	0.7143353	0.64006937	309
rpsM	ribosomal protein S13	0.73769516	0.700627	0.63180536	366
rplM	ribosomal protein L13	0.7012361	0.7173054	0.63109225	438
rplT	ribosomal protein L20	0.7059461	0.7043889	0.6296289	360
rpsG	ribosomal protein S7 (BS7)	0.69414663	0.6983668	0.62722415	471
rpmA	ribosomal protein L27 (BL24)	0.7617893	0.68935204	0.62463117	285
rplL	ribosomal protein L12 (BL9)	0.7505788	0.75000113	0.70929384	372
rpsI	ribosomal protein S9	0.6746212	0.6966346	0.62250024	393
rplB	ribosomal protein L2 (BL2)	0.68995285	0.7142757	0.6219602	834
rpsD	ribosomal protein S4 (BS4)	0.66974896	0.6913568	0.61133987	603
rplX	ribosomal protein L24 (BL23) (histone-like protein HPB12)	0.69466513	0.6841065	0.60728997	312
rpsF	ribosomal protein S6 (BS9)	0.6685666	0.670196	0.59957814	288
rpsE	ribosomal protein S5	0.66058165	0.6591603	0.5988015	501
rplD	ribosomal protein L4	0.67654395	0.6655937	0.59404176	624
rplJ	ribosomal protein L10 (BL5)	0.6426427	0.6662391	0.59200644	501
rplP	ribosomal protein L16	0.69511014	0.6698322	0.5890895	435
rplN	ribosomal protein L14	0.69714874	0.6653969	0.58674556	369

Table 8. Reference set of 43 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; gCAI = computed with the algorithm in 15 iterations; 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	gCAI	Len
tufA	protein chain elongation factor EF-Tu	0.81218296	0.79347014	0.75036675	1185
tufB	protein chain elongation factor EF-Tu	0.7899474	0.7695716	0.72075397	1185
tsf	protein chain elongation factor EF-Ts	0.7634853	0.7492837	0.71321905	852
fusA	GTP-binding protein chain elongation factor EF-G	0.7482026	0.7281189	0.66172844	2115
mopA	chaperonin GroEL	0.7897918	0.7831986	0.7468645	1647
dnaK	heat shock protein DnaK	0.7178141	0.70362157	0.64333135	1917
cspA	cold shock protein 7.4	0.7569456	0.75185895	0.71599495	213
tig	trigger factor	0.7310184	0.7203704	0.6729257	1299
ompA	outer membrane protein	0.7801367	0.7642526	0.71205974	1041
ompX	outer membrane protein	0.7242723	0.71460575	0.6599186	516
ompC	outer membrane protein	0.812891	0.7957519	0.74745786	1104
lpp	murein lipoprotein	0.80520666	0.79619217	0.76923805	237
pal	peptidoglycan-associated lipoprotein	0.6727051	0.6656418	0.60742617	522
yaiU	putative flagellin structural protein	0.70819515	0.703465	0.6458635	1404
yfiD	putative formate acetyltransferase	0.67965895	0.67101467	0.61617833	384
eno	diadenosine tetraphosphatase	0.8348635	0.82159644	0.7851021	1299
tpiA	triosephosphate isomerase	0.73652947	0.72718644	0.66838896	768
pgk	phosphoglycerate kinase	0.72215086	0.7125015	0.6625036	1164
gapA	glyceraldehyde-3-phosphate dehydrogenase A	0.8279393	0.8149085	0.7754178	996
fba	fructose-bisphosphate aldolase class II	0.7683944	0.7610465	0.70846814	1080
pykF	pyruvate kinase I	0.7016575	0.6887916	0.62648696	1413
pflB	formate acetyltransferase 1	0.77904093	0.7667218	0.7108569	2283
ahpC	alkyl hydroperoxide reductase C22 subunit	0.783308	0.77746105	0.7265764	564
sodA	superoxide dismutase SodA	0.7065987	0.70124394	0.6502629	621
tktA	transketolase 1/2 isozyme	0.68566346	0.68041885	0.6122929	1992
rpoC	RNA polymerase beta prime subunit	0.69955987	0.6894705	0.6256046	4224
rpsI	30S ribosomal subunit protein S9	0.7563791	0.74349326	0.6916582	393
rpsA	30S ribosomal subunit protein S1	0.7740142	0.757111	0.71050775	1674
rpsB	30S ribosomal subunit protein S2	0.76608026	0.75217986	0.7085285	726
rpsC	30S ribosomal subunit protein S3	0.7213994	0.7171848	0.66578275	702
rpsU	30S ribosomal subunit protein S21	0.68349123	0.67777705	0.6431132	216
rplA	50S ribosomal subunit protein L1	0.75861174	0.7458433	0.7101479	705
rplY	50S ribosomal subunit protein L25	0.72813624	0.7222367	0.6834064	285
rplI	50S ribosomal subunit protein L9	0.70551	0.7004429	0.65935963	450
rplL	50S ribosomal subunit protein L7/L12	0.81440103	0.8045188	0.78317964	366
rplC	50S ribosomal subunit protein L3	0.6988004	0.6896239	0.6389732	630
rpmE	50S ribosomal subunit protein L31	0.6974921	0.69375414	0.6343497	213
rplB	50S ribosomal subunit protein L2	0.7024668	0.6895137	0.63052315	822
rplK	50S ribosomal subunit protein L11	0.6863151	0.6745547	0.6287197	429
rpmI	50S ribosomal subunit protein A	0.66723424	0.6622692	0.62579	198
rpmA	50S ribosomal subunit protein L27	0.70560354	0.67746854	0.62051374	258
rplD	50S ribosomal subunit protein L4, regulates expression of S10 operon	0.6872925	0.66999024	0.6178337	606

Table 7. Reference set of 44 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; the "gCAI" value is calculated in 15 iterations; "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	gCAI	Len
TDH1	Glyceraldehyde-3-phosphate dehydrogenase 1	0.84769833	0.85030264	0.8127145	999
TDH2	glyceraldehyde 3-phosphate dehydrogenase	0.8951536	0.8981491	0.8730602	999
TDH3	Glyceraldehyde-3-phosphate dehydrogenase 3	0.9137378	0.91561896	0.8930528	999
PDC1	pyruvate decarboxylase	0.8994493	0.9010076	0.8702157	1692
ENO1	enolase I	0.86449337	0.8649675	0.83544266	1314
ENO2	enolase	0.88618064	0.8856355	0.8593889	1314
FBA1	aldolase	0.8612911	0.86331147	0.8182507	1080
TPI1	triosephosphate isomerase	0.8047238	0.81116617	0.7714424	747
GPM1	converts 3-phosphoglycerate to 2-phosphoglycerate in glycolysis	0.801199	0.7989722	0.7518571	744
ADH1	Alcohol dehydrogenase	0.80602115	0.8002969	0.74935853	1047
ILV5	ketol-acid reducto-isomerase	0.7990355	0.803838	0.75177014	1188
CDC19	pyruvate kinase	0.88755256	0.8884879	0.85680336	1503
TEF2	translational elongation factor EF-1 α	0.87031275	0.8693718	0.8239393	1377
TEF1	translational elongation factor EF-1 α	0.865811	0.8653603	0.81933904	1377
EFT1	translation elongation factor 2 (EF-2)	0.8056308	0.80909216	0.76125926	2529
EFT2	translation elongation factor 2 (EF-2)	0.80214024	0.80541897	0.75631016	2529
YEF3	translation elongation factor eEF3	0.78085816	0.7831822	0.7184399	3135
HYP2	Translation initiation factor eIF-5A	0.79853266	0.8058099	0.758234	474
SSB1	heat shock protein of HSP70 family	0.81719404	0.82347804	0.77961385	1842
SSA2	member of 70 kDa heat shock protein family	0.79906124	0.8019022	0.75686127	1920
PGK1	3-phosphoglycerate kinase	0.80878544	0.80267113	0.74991524	1251
CCW12	Similarity with FL01, flocculation	0.84228027	0.84676874	0.8047048	402
ASC1	40S small subunit ribosomal protein	0.77717847	0.7812206	0.7199785	960
RPL39	ribosomal protein L39	0.8435371	0.83970976	0.81139606	156
RPL30	ribosomal protein L30	0.8301169	0.8311089	0.8055944	318
RPL22A	ribosomal protein L22	0.82883966	0.83765775	0.7993247	366
RPS6B	ribosomal protein S6	0.8310364	0.83364505	0.7939496	711
RPS12	ribosomal protein S12	0.82209843	0.8275626	0.7936337	432
RPL43A	ribosomal protein L37a	0.8333912	0.8344563	0.7912244	279
RPL8B	ribosomal protein L7a	0.8365352	0.8340938	0.79063654	771
RPL5	ribosomal protein L5; for assembly of stable 60S rib. subun.	0.81914777	0.8248273	0.78788507	894
RPS5	ribosomal protein S5	0.8240064	0.8255974	0.7870322	678
RPL8A	ribosomal protein L7a	0.82941675	0.82414716	0.78256255	771
RPL1B	60S large subunit ribosomal protein	0.81982213	0.82778186	0.7792641	654
RPS6A	ribosomal protein S6	0.8225214	0.82250047	0.7756098	711
RPL4B	ribosomal protein L4A	0.80601734	0.8124044	0.76995146	1089
RPL10	acidic ribosomal protein L10.e	0.8141283	0.8233417	0.76930636	666
RPL3	ribosomal protein L3	0.82569766	0.8239667	0.7669779	1164
RPL1A	ribosomal protein L10a	0.80849254	0.8168908	0.76649565	654
RPS20	ribosomal protein S20	0.80002564	0.8118208	0.76593626	366
RPL18A	ribosomal protein L18	0.7941487	0.8010349	0.7631792	561
RPS24A	ribosomal protein S24	0.7896517	0.79360914	0.7616996	408
RPL17A	ribosomal protein L17	0.791737	0.7950371	0.7591775	555
RPL4A	ribosomal protein L4	0.7969803	0.80131054	0.75761247	1089
RPS22A	ribosomal protein S15a	0.78869784	0.79102826	0.75121105	393
RPS10A	ribosomal protein S10	0.8076113	0.80458766	0.7497459	318
RPS14A	ribosomal protein S14	0.78170025	0.7881132	0.74724674	414
RPS9B	ribosomal protein S9	0.7925195	0.7944186	0.74512553	588
RPL32	ribosomal protein L32	0.7930637	0.7941196	0.74458504	393
RPP0	acidic ribosomal protein L10.e	0.7862089	0.7907318	0.74283695	939
RPP2A	acidic ribosomal protein P2. β	0.77294767	0.7754243	0.74126875	321
RPS3	ribosomal protein S3	0.78942317	0.7856124	0.74115723	723
RPS17A	ribosomal protein S17	0.78502053	0.78344977	0.73347527	411
RPS2	ribosomal protein S2	0.79019827	0.7927135	0.7332593	765
RPL36B	ribosomal protein L36	0.76923275	0.76645136	0.73222715	303
RPS19A	ribosomal protein S19	0.7741291	0.7708095	0.7313616	435
RPS23B	ribosomal protein S23 and E. coli S12	0.768031	0.7739426	0.73034954	438
RPS31	ribosomal protein S27a	0.7877448	0.78993887	0.72774374	459
RPL15A	ribosomal protein L15	0.7708423	0.77133524	0.7248617	615
RPS26A	ribosomal protein S26	0.75402683	0.761466	0.7225349	360
RPL2A	ribosomal protein	0.76163936	0.7656473	0.72191304	765

Table 9. 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; gCAI = computed with the algorithm in 15 iterations; 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	gCAI	Con	Exp	Len
T28C6.6	CE03749	cuticle collagen	0.7993916	0.8159407	0.7608275	PC		900
T28C6.4	CE03749	cuticle collagen	0.7993916	0.8159407	0.7608275	C		900
Y41E3.2	CE18370	cuticle collagen 1	0.7655781	0.76761854	0.6986232	PC		870
W03G11.1	CE26826	cuticle collagen	0.679989	0.7241507	0.6462018	PC		840
sqt-1		cuticle collagen SQT-1	0.6830913	0.71224433	0.6353661			975
C34H4.4	CE27821	cuticular collagen	0.81204915	0.8110274	0.75991553	PC		711
F27C1.8	CE09720	cuticular collagen	0.766677	0.7923744	0.7293668	PC		855
F30B5.1	CE04462	cuticular collagen	0.77667356	0.7835335	0.7177716	C		909
F26F12.1	CE07146	cuticular collagen	0.7561981	0.7617512	0.6921289	PC		870
ZC373.7	CE02380	cuticular collagen	0.70891196	0.7417199	0.6669354	PC		894
B0222.6	CE30585	collagen	0.8255235	0.8511477	0.8005858	PC	NE	852
F55C10.3	CE11182	collagen	0.71783507	0.73821396	0.6596515	P		801
T06E4.6	CE06372	collagen	0.6887894	0.7241842	0.6527216	P		873
C44C10.1	CE05408	collagen	0.6944948	0.7167036	0.6445038	PC		903
F55C10.2	CE05952	collagen	0.696306	0.72491425	0.64805573	C		900
C09G5.5	CE01485	collagen	0.7296231	0.7583281	0.6894046	PC		954
F08G5.4	CE05574	collagen	0.71982646	0.75062704	0.6821709	PC		900
F38A3.1	CE02213	collagen	0.70065933	0.7373799	0.6626436	PC		921
F23H12.4	CE05707	collagen	0.7164265	0.7366751	0.6617057	PC		906
C53B4.5	CE03091	collagen	0.7175056	0.7372882	0.6615382	PC		870
C29F4.1	CE03038	collagen	0.79276764	0.8091042	0.7533284	PC		900
F38A3.2	CE02214	collagen	0.76600504	0.78904593	0.726436	C		921
F11G11.10	CE31783	collagen	0.75049764	0.77706945	0.7110588	C	NE	852
F57B1.3	CE11290	collagen	0.74670154	0.7757267	0.7122179	PC		942
F11G11.12	CE09354	collagen	0.7436009	0.770927	0.70627195	PC	NE	858
C09G5.4	CE01484	collagen	0.69363374	0.71582556	0.6414115	PC		972
F54D1.2	CE05937	collagen	0.666996	0.70720524	0.63500047	PC		885
T15B7.4	CE13654	collagen	0.6700034	0.70023936	0.62511575	PC	NE	978
T06E4.4	CE06370	collagen	0.6695214	0.7007443	0.6241404	PC		873
T11F9.9	CE06421	collagen	0.6612906	0.69793123	0.6172368	P		900
W05B2.1	CE20144	collagen	0.7911253	0.8134583	0.7582315	C		915
B0222.7	CE06698	collagen	0.7795063	0.8027505	0.73923653	PC	NE	885
W05B2.6	CE20148	collagen	0.77408165	0.79630834	0.73411095	PC		915
col-12		collagen	0.6677796	0.7126159	0.6418425			951
F53F1.4	CE10938	cuticlin	0.7858718	0.8054621	0.75262064	C		318
F53F1.5	CE10940	cuticlin	0.7699064	0.8001557	0.7455146	C		648
F41F3.3	CE07217	cuticlin	0.7918024	0.7822662	0.7266986	C		471
T04C12.6	CE13148	actin	0.78852284	0.79210633	0.72923744	C		1131
act-1		actin	0.78852284	0.79210633	0.72923744			1131
act-4		actin	0.7887807	0.7871011	0.72611415			1131
T04C12.5	CE13150	actin	0.710738	0.7250126	0.65228915	C		1131
Y75B12B.2	CE20371	Peptidyl-prolyl cis-trans isomerases	0.75694305	0.7743963	0.7145612	C		516
Y75B12B.5	CE20374	Peptidyl-prolyl cis-trans isomerases	0.7424295	0.7509994	0.6805518	C		522
C07A12.4	CE03972	protein disulfide isomerase	0.8130993	0.83383375	0.77327424	C		1482
sip-1		heat shock hsp20 proteins	0.684132	0.6923746	0.6146739			480
C47E8.5	CE05441	heat shock protein (HSP90)	0.7435628	0.7769464	0.7144739	C		2109
hsp-1		heat shock 70kd protein A	0.72286123	0.75082576	0.6860208			1923
hsp-3		heat shock protein	0.7095725	0.73370206	0.65554374			1062
myo-1		myosine	0.6593437	0.6964941	0.6244561			3246
ost-1		osteonectin	0.7251205	0.747736	0.6607431			795
C47B2.3		α -tubulin	0.7403996	0.7561489	0.68163			1347
T22E5.5	CE04994	troponin T-like protein	0.72671175	0.75012887	0.6797876	C		1218

Table 10 (51 genes out of 169). Reference set of 169 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	gCAI	Con	Exp	Len
F55D10.2	CE02777	Ribosomal protein L23	0.7931173	0.8139563	0.7650799	C		444
E04A4.8	CE21392	ribosomal protein	0.82837003	0.8454685	0.7957217	C	NE	336
F56F3.5	CE00664	Ribosomal protein S3a (human) homolog.	0.7817147	0.802747	0.7446678	C		774
B0393.1	CE00854	40S ribosomal protein	0.77228665	0.7993645	0.73869133	C		831
D1007.12	CE09047	60S ribosomal protein L24	0.7698794	0.79286695	0.7365045	C		480
F53A3.3	CE10884	40S ribosomal protein	0.79215145	0.80140495	0.73501724	C		393
K11H12.2	CE12148	60S ribosomal protein L15	0.75319636	0.7777614	0.7106428	C		615
C23G10.3	CE01810	Ribosomal protein S3	0.74707514	0.77017266	0.70953465	C		744
F07D10.1	CE07033	ribosomal protein	0.7700089	0.7698313	0.7084613	C		591
F13B10.2	CE05598	60S ribosomal protein L3	0.7434626	0.77405095	0.70132756	C		606
B0513.3	CE15559	60S ribosomal protein L29	0.7514963	0.760861	0.6998169	C		189
B0041.4	CE07669	ribosomal protein L1	0.72817355	0.7564101	0.6894746	C		1038
F28C6.7a	CE03278	ribosomal protein L26 like	0.7311577	0.7518051	0.6854304	C		429
F39B2.6	CE16012	40S ribosomal protein S26	0.7239612	0.7563946	0.6872341	C		354
ZC434.2	CE06577	40S ribosomal protein S7	0.71903104	0.74631745	0.6795441	C		585
C09H10.2	CE02131	60S ribosomal protein	0.718919	0.737804	0.6784309	C		318
F10E7.7	CE04362	60s ribosomal protein L35A	0.70891887	0.734615	0.6757259	C		375
T05E11.1	CE06360	40S ribosomal protein S5	0.7119159	0.74543107	0.6752486	C		633
D1007.6	CE09041	40S ribosomal protein S10	0.7088387	0.73145103	0.6644497	C		450
F13B10.2	CE05598	60S ribosomal protein L3	0.71965605	0.72689503	0.656473	C		600
F53G12.10	CE11024	ribosomal protein	0.7109244	0.7273467	0.65848243	C		735
F42C5.8	CE04561	40S ribosomal protein S8	0.7033947	0.72536606	0.6515279	C		627
F37C12.9	CE00821	Ribosomal protein S14	0.68226635	0.7146262	0.6459627	C	NE	459
Y57G11C.16	CE14956	ribosomal protein S13	0.7076305	0.7173757	0.65320814	C		465
Y45F10D.12	CE16650	Eukaryotic ribosomal protein L18	0.7105242	0.7233832	0.65312356	C		567
F54C9.5	CE02255	60S ribosomal protein L5	0.69843596	0.7179434	0.64325774	C		882
C32E8.2	CE08526	ribosomal protein L13	0.6914838	0.71108633	0.64039814	C		624
R11D1.8	CE06313	ribosomal protein L28 like	0.6700374	0.7080754	0.6368416	C		381
C26F1.4	CE06878	ribosomal protein/ubiquitin-like protein	0.6815946	0.70534325	0.6358191	C		393
T24B8.1	CE03709	60S ribosomal protein L32	0.6939911	0.7052838	0.635158	C		405
B0250.1	CE18478	Ribosomal Proteins L2	0.6998294	0.7082962	0.63477916	C		783
F40F11.1	CE05860	ribosomal protein S11	0.68963724	0.7146166	0.630872	C		468
R151.3	CE00744	Ribosomal protein ML16	0.6661125	0.7032525	0.6279336	C		654
F10B5.1	CE01543	ribosomal protein L10 (QM protein)	0.699068	0.714598	0.6266797	C		645
F36A2.6	CE09945	40S ribosomal protein S15	0.6799013	0.69294155	0.61598754	C		456
C49H3.11	CE04237	ribosomal protein rps2	0.72226983	0.72841644	0.6599411	C		819
F54E7.2	CE26896	orthologous to <i>S. cerevisiae</i> rib.prot. rps12	0.71584463	0.74487066	0.6643111	C		438
JC8.3		orthologous to <i>S. cerevisiae</i> rib. prot. rpl23	0.7294716	0.7581929	0.698507	C		498
B0336.10	CE00778	orthologous to <i>S. cerevisiae</i> rib. prot. rpl23	0.75076354	0.77271277	0.72025096	C	NE	423
Y106G6H.3	CE20413	structural protein of ribosome rpl-30	0.64869124	0.6947521	0.62922305	PC		342
C06B8.8	CE20485	structural protein of ribosome rpl-38	0.68528014	0.6999021	0.6184305	C		213
C04F12.4	CE19677	structural protein of ribosome rpl-14	0.66441625	0.692433	0.6252308	C		408
R09B3.2	CE16307	RNA recogn. motif (aka RRM/RBD/RNP dom.)	0.69566435	0.72037375	0.66532373	C		252
R09B3.3	CE16308	RNA recogn. motif (aka RRM/RBD/RNP dom.)	0.6900253	0.719639	0.64780533	C		258
R06C1.4	CE18119	RNA recogn. motif. (aka RRM/RBD/RNP dom.)	0.74324733	0.7624747	0.71824855	C		255
R03G5.1	CE01270	elongation factor EF-1 α	0.8211111	0.84100306	0.787374	C		1392
eft-2		elong. factor Tu fam. (ATP/GTP binding P-loop)	0.69113755	0.71332	0.6386557			2559
eft-3		elongation factor 1- α	0.81235236	0.8357108	0.78254473			1392
T05G5.10	CE00503	Initiation factor 5A	0.6567467	0.69578725	0.61964226	C		486
F54C9.1	CE02249	initiation factor 5A	0.7431964	0.7613361	0.68646985	C		486
ubq-2		UBQ-2 ubiquitin; 60S rib.prot. L40	0.694645	0.7101069	0.63295406			387
E04A4.7	CE16968	cytochrome C	0.7066116	0.7399022	0.67512184	C		336
F26E4.9	CE09693	cytochrome C oxidase	0.6629373	0.70047134	0.6256358	C		399
C01G6.1a	CE00863	glycerol uptake facilitator protein	0.71093005	0.7367888	0.667226	C		873
T27E9.1	CE14263	ADP/ATP carrier protein	0.73368216	0.76894915	0.705634	C		903
unc-15		enzyme - carbohydrate metabolism	0.73995197	0.766195	0.705222			2274
unc-54		non-muscle myosin, ATP binding	0.73974633	0.7426133	0.665905			2852
H28O16.1	CE18826	ATP synth. α/β subun., ATP synth. α -chain, C-term.	0.7258207	0.74966186	0.68493444	C		1617
F32D1.2	CE09866	ATP synthase ϵ -chain	0.66556	0.6940483	0.6256993	C		165
cpr-4		cathepsin B-like cysteine proteinase	0.73206514	0.77176636	0.6979663			1008

Table 10 (continued, 60 genes out of 169). Reference set of 169 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&a.l.	CAI	gCAI	Con	Exp	Len
F20H11.3	CE09512	malate dehydrogenase	0.652351	0.6947127	0.62921953	C		1026
ZK829.4	CE06652	glutamate dehydrogenase	0.6756591	0.7084176	0.6340552	C		1611
gpd-2		glyceraldehyde 3-phosphate dehydrogenase	0.7602503	0.7969548	0.73505664			1026
gpd-3		glyceraldehyde 3-phosphate dehydrogenase	0.7678847	0.8043856	0.74229157			1026
F13D12.4	CE02183	methylmalonate-semialdehyde dehydrogenase	0.67485	0.70075756	0.62011707	C		1572
F46E10.10	CE20820	lactate dehydrogenase	0.66114396	0.69207495	0.6151204	C		1011
H02I12.6	CE05165	Core histones H2A, H2B, H3 and H4	0.68430364	0.7136602	0.6464371	PC		372
Y49E10.6	CE22223	Core histone H2A/H2B/H3/H4	0.69397634	0.7054898	0.6422423	C		411
H02I12.7	CE04501	Core histones H2A, H2B, H3 and H4	0.6723196	0.702381	0.6417951	PC		384
F08G2.1	CE07075	Core histones H2A, H2B, H3 and H4	0.6553546	0.69864017	0.6261421	PC		369
F22F1.1	CE04443	histone H1	0.66658175	0.7084917	0.6503549	C		627
his-24		histone H1	0.74666816	0.7800629	0.7353782			627
F54E12.5	CE04501	histone H2A	0.6723196	0.702381	0.6417951	P		384
B0035.7	CE04501	histone H2A	0.6723196	0.702381	0.6417951	PC		384
F45F2.12	CE10538	histone H2B	0.6617916	0.70071226	0.64090383	P	NE	372
B0035.8	CE05165	histone H2B	0.6741838	0.70409304	0.63526255	C		372
F54E12.4	CE05165	histone H2B	0.68430364	0.7136602	0.6464371	P		372
ZK131.9	CE07075	histone H2B	0.6553546	0.69864017	0.6261421	P		369
ZK131.5	CE07075	histone H2B	0.6553546	0.69864017	0.6261421	PC		369
his-11		histone H2B	0.71548045	0.75535274	0.6989031			369
T10C6.13	CE03253	histone H3	0.6816341	0.7205982	0.6501348	PC		411
his-9		histone H3	0.6699781	0.7055368	0.63480943			411
F02A9.2	CE00133	<i>O. volvulus</i> 20Kd antigenic peptide	0.8208756	0.84626156	0.797979	C		549
C06A8.3	CE02454	17k antigen (<i>O. volvulus</i>)	0.7138214	0.7515148	0.69157046	C		456
F25H2.11	CE09656	TCTP protein	0.7404227	0.7717513	0.71165365	C		546
F25H2.5	CE09650	nucleoside diphosphate kinase	0.7602652	0.7783908	0.711409	C		462
F46H5.3	CE04589	arginine kinase	0.7629622	0.7806214	0.7058248	C		1134
C01G6.1b	CE19664	similarity to transport membrane protein	0.7126592	0.7389466	0.6692718	C		858
F21F8.7	CE09542	protease	0.7050553	0.74223757	0.6670415	C		1170
H22K11.1	CE19495	aspartyl protease	0.70120317	0.7344943	0.65467566	C		1197
K04D7.1	CE06090	guanine nucleotide-binding protein	0.707762	0.7343344	0.65309066	C		978
F52D10.3		similarity to 14-3-3 family of proteins	0.7182883	0.74018633	0.66703767			747
Y57G11C.15	CE14954	protein transport protein SEC61 α -subunit	0.7021698	0.72011006	0.6429967	PC		1422
ZK180.5		weak simil. to CCAAT/enhancer binding prot. β	0.6899541	0.7099849	0.64111745		NE	1629
C49F5.1	CE08852	s-adenosylmethionine synthetase	0.67462367	0.70999783	0.63939303	C		1212
Y38A10A.5	CE21562	calreticulin precursor	0.6950712	0.7211511	0.6360596	C		1188
F25H2.10	CE09655	deoxyribonuclease	0.72335505	0.74507	0.682776	C		939
vit-1		pseudogene 95% homologous to vit-2	0.6659802	0.7007626	0.6153985			4851
vit-2		vitellogenin	0.7358966	0.7646283	0.69552916			4842
vit-5		structural gene for vitellogenin yp170A	0.68738854	0.7269998	0.65246606			4812
vit-6		vitellogenin	0.7263971	0.7579766	0.6825292			4956
C56C10.8	CE02573	Transcription factor BTF3 (human)	0.65368927	0.6925	0.6242902	C		486
K02F2.2	CE17154	S-adenosylhomocysteine hydrolase	0.7793253	0.80826205	0.75087124	C		1314
Y47D3B.6	CE24344	similarity with PFAM domain	0.66513723	0.6999882	0.61349916	PC		852
C27A2.2		hypothetical protein	0.755627	0.7860153	0.7291372			393
F59F4.2	CE11548	unknown function	0.72338307	0.7531682	0.6826055	C		198
F09F7.2		unknown function	0.72716236	0.737876	0.6751239			462
C01B10.5	CE06736	unknown function	0.661388	0.7065796	0.6647513			189
C02E7.6	CE07843	unknown function	0.71657014	0.73726755	0.68241495	C		213
F20A1.1	CE07091	unknown function	0.6376904	0.6884248	0.6263674	C	NE	342
F44E5.1	CE18676	unknown function	0.6595811	0.69055593	0.61947584	C		237
F47B7.1	CE02743	unknown function	0.69552004	0.71947694	0.64245343	C	NE	180
F11C7.5	CE17657	unknown function	0.67648286	0.70182306	0.6142195	C		570
R102.2	CE23912	unknown function	0.7092644	0.73547655	0.65677124	C		342
B0222.8	CE06699	unknown function	0.8421129	0.87207395	0.82230556	C		885
F02A9.3	CE00134	unknown function	0.7918183	0.8224969	0.7793565	C		549
C16A3.8	CE27691	unknown function	0.75158864	0.763057	0.6960954	PC		456
F46F2.3	CE05880	unknown function	0.7466418	0.75087714	0.6959811	C		414

Table 10 (continued, 58 genes out of 169). Reference set of 169 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&a.l. = 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	gCAI	Len
Hsp83	Heat shock protein 83	0.7625245	0.71697533	0.65732574	2154
Hsc70-4	Heat shock protein cognate 4	0.7481074	0.7029956	0.64601916	1956
Hsp60	Heat shock protein 60	0.72055304	0.6957708	0.6411236	1722
Pdi	Protein disulfide isomerase	0.69148105	0.67348254	0.6096634	1491
CG2852	peptidyl-prolyl cis-trans isomerase	0.83065456	0.79658586	0.75418293	363
CG2852	peptidyl-prolyl cis-trans isomerase	0.7590536	0.7288332	0.6673145	618
CG5520	chaperone	0.67736846	0.66977316	0.60009676	2364
Cyt-c-p	Cytochrome c proximal	0.7312841	0.6997262	0.64745617	327
CG3762	hydrogen-transporting two-sector ATPase	0.7172716	0.6660187	0.59567606	1845
SdhB	Succinate dehydrogenase B	0.6790003	0.70606625	0.63140124	894
Vha55	Vacuolar H[+]-ATPase 55kD B subunit	0.69689727	0.6654428	0.59519327	1473
CG14028	cytochrome c oxidase respiratory chain complex IV	0.74593616	0.71515346	0.6673668	240
ATPsyn- γ	ATP synthase- γ chain	0.66725314	0.65271825	0.5872132	894
CG17280	cytochrome c oxidase respiratory chain complex IV	0.7265877	0.7288774	0.66040677	330
ATPsyn-d	ATP synthase, subunit d	0.6784137	0.67614955	0.6129384	537
CG3861	citrate (S1)-synthase mitochondrial matrix	0.6728218	0.6670812	0.5989491	1395
Ald	Aldolase	0.79484934	0.7437344	0.69133466	1086
Eno	Enolase	0.80140764	0.7809171	0.7300683	1302
Tpi	Triose phosphate isomerase	0.76233506	0.7718421	0.709473	744
Gdh	Glutamate dehydrogenase	0.7243763	0.7155702	0.64797544	1695
CG5548	NADH dehydrogenase	0.644987	0.6791958	0.60762066	354
CG7998	malate dehydrogenase 1 mitochondrial matrix	0.72942233	0.7055823	0.6524246	1011
CG15093	3-hydroxyisobutyrate dehydrogenase mitochondrion	0.6764571	0.6618336	0.5921798	975
CG11151	estradiol 17 β -dehydrogenase	0.6993486	0.69393873	0.6422715	348
Adh	Alcohol dehydrogenase	0.69577736	0.66349417	0.6052148	771
Gapdh1	Glyceraldehyde 3 phosphate dehydrogenase 1	0.722122	0.6866197	0.62490505	996
Act57B	Actin 57B	0.78972256	0.7356825	0.67863405	1131
Act5C	Actin 5C	0.7380564	0.69550616	0.63042563	1131
Act79B	Actin 79B	0.7111434	0.69392383	0.6202902	1131
Act88F	Actin 88F	0.70571494	0.67592365	0.5984086	1131
α Tub84B	α -Tubulin at 84B	0.7581589	0.7104684	0.6482581	1353
α Tub84B	α -Tubulin at 84B	0.7527154	0.7080009	0.6452747	1287
β Tub56D	β -Tubulin at 56D	0.67546123	0.6696437	0.5952855	1344
Chd64	actin binding	0.75223005	0.7596788	0.7058922	528
Tm2	Tropomyosin 2	0.70480156	0.70789796	0.64848393	855
Mlc-c	Myosin light chain cytoplasmic	0.6979156	0.68278855	0.6233812	444
Prm	Paramyosin	0.67170477	0.6767034	0.60685664	1434
Prm	Paramyosin	0.67063236	0.66140705	0.59611547	2640
Cdle2	Cytoplasmic dynein light chain 2	0.7623953	0.75094664	0.7042127	270
Lcp65Ac	cuticle (similarity)	0.68017715	0.6481334	0.5927398	330
Ccp84Ad	cuticle (similarity)	0.6812113	0.6502954	0.58921146	600

Table 11 (40 genes out of 128). Reference set of 128 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; gCAI = computed with the algorithm in 15 iterations; 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	gCAI	Len
Ef1 α 48D	Elongation factor 1 α 48D	0.7606497	0.6960081	0.63635534	1392
Ef1 γ	involved in translational elongation	0.73996735	0.7404281	0.68151647	1293
CG9282	involved protein biosynthesis	0.65348357	0.65852153	0.5879676	468
CG9091	involved in protein biosynthesis	0.714825	0.7006865	0.6300774	282
CG8332	involved in protein biosynthesis	0.7006849	0.67248255	0.60591996	447
CG1883	involved in protein biosynthesis	0.7332427	0.70562696	0.65158117	585
CG3751	involved in protein biosynthesis	0.72553444	0.6982998	0.6426775	396
CG1475	involved in protein biosynthesis	0.7145987	0.69965667	0.63998204	354
CG4759	involved in protein biosynthesis	0.6954164	0.68363124	0.6169676	408
CG4111	involved in protein biosynthesis	0.6779047	0.6788727	0.62619615	366
CG6846	involved in protein biosynthesis	0.74522436	0.73794496	0.6881972	450
RpL11	involved in protein biosynthesis	0.66750324	0.65215796	0.5855542	555
RpS12	Ribosomal protein S12	0.68146485	0.65224874	0.5875348	420
RpS17	Ribosomal protein S17	0.7064104	0.67066395	0.60383016	396
RpL32	Ribosomal protein L32	0.6979241	0.6863202	0.61301255	405
RpS18	Ribosomal protein S18	0.79379886	0.7528799	0.6988274	459
RpS6	Ribosomal protein S6	0.7497512	0.69129467	0.6343729	654
RpL40	Ribosomal protein L40	0.71698433	0.68216115	0.62236077	387
RpL18A	Ribosomal protein L18A	0.70980203	0.69829804	0.62425673	534
RpL8	Ribosomal protein L8	0.77507406	0.72210264	0.6604139	771
RpL7A	Ribosomal protein L7A	0.7170834	0.6994674	0.63667285	816
RpP1	Ribosomal protein P1	0.6907852	0.6582836	0.61115634	342
RpL3	Ribosomal protein L3	0.7662649	0.7412646	0.6814405	1212
RpS9	Ribosomal protein S9	0.7186592	0.67797536	0.61424685	588
RpL36	Ribosomal protein L36	0.6686283	0.6634981	0.60073954	348
RpL13	Ribosomal protein L13	0.69913507	0.6851982	0.6199556	657
RpP2	Ribosomal protein P2	0.76689297	0.720888	0.68070024	339
RpL1	Ribosomal protein L1	0.715556	0.6872822	0.61701435	1206
RpS26	Ribosomal protein S26	0.6930774	0.6944534	0.6287494	345
CG11522	large ribosomal subunit (component)	0.7374945	0.7300337	0.67662203	789
CG11522	large ribosomal subunit (component)	0.73222184	0.7277298	0.6734491	732

Table 11 (continued, 30 genes out of 128). Reference set of 128 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; gCAI = computed with the algorithm in 15 iterations; 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	gCAI	Len
CG12628	glutathione transferase	0.7037151	0.68765324	0.6166774	366
Yp1	yolk protein vitellogenesis	0.6936118	0.6703773	0.6066335	1320
CG14745	peptidoglycan recogn. integral plasma membrane prot.	0.6652215	0.66334283	0.5920152	555
Fer1HCH	Ferritin 1 heavy chain homologue	0.7107032	0.7002529	0.6386484	618
DebB	Developmental embryonic B	0.6486722	0.65198004	0.59451556	255
m4	E(spl) region transcript m4	0.66728616	0.65737414	0.594107	459
Vm26Aa	Vitelline membrane 26Aa	0.7161637	0.67874604	0.6211306	426
smt3	involved in protein-nucleus import	0.6987937	0.65699106	0.5970316	273
γ Try	γ Trypsin	0.6910861	0.64975595	0.5940643	426
blw	bellwether	0.7923425	0.73553526	0.68543273	1659
Sod2	Superoxide dismutase 2 (Mn)	0.6441058	0.66231954	0.58831686	654
CG18111	odorant binding	0.8147616	0.8208617	0.76990163	354
CG18730	α -amylase	0.7404946	0.7379515	0.66970146	1482
CG14746	peptidoglycan recognition integral plasma membrane	0.7422852	0.7359498	0.6684485	558
CG8577	peptidoglycan recognition integral plasma membrane	0.72720784	0.71723723	0.64755803	558
CG9480	glycogenin glucosyltransferase	0.6793733	0.65876216	0.585713	1275
Dbi	Diazepam-binding inhibitor	0.70571494	0.67592365	0.5984086	1131
CG8629	acyl-CoA homeostasis	0.70644367	0.6888298	0.63689905	255
CG8628	acyl-CoA homeostasis	0.65988654	0.6436429	0.5927783	255
CG15829	acyl-CoA homeostasis	0.7715402	0.7463332	0.6995288	249
CG5804	acyl-CoA homeostasis	0.77383447	0.7395119	0.7010428	249
LysC	Lysozyme C	0.69248	0.66369754	0.60764396	129
GlyP	Glycogen phosphorylase	0.67833155	0.67724174	0.60349756	2535
Ser99Da	Serine protease 1	0.7233682	0.6986089	0.62749064	798
Ser99Db	Serine protease 2	0.704828	0.6773687	0.60314393	759
GstD1	Glutathione S transferase D1	0.7206267	0.71448857	0.6518124	630
ImpL3	Ecdysone-inducible gene L3	0.6787649	0.666634	0.5949092	999
oho23B	overgrown hematopoietic organs at 23B	0.7672127	0.7491063	0.6984381	252
Argk	Arginine kinase	0.8015121	0.76288444	0.71334714	1071
Lsp1 β	Larval serum protein 1 β	0.83072245	0.8292647	0.78096724	2370
Mp20	Muscle protein	0.7791874	0.7471312	0.70314837	555
Amy-d	Amylase distal	0.7380434	0.73737264	0.6704118	1482
Ser4	Serine protease 4	0.77213806	0.7306876	0.6696332	801
Rack1	Receptor of activated protein kinase C 1	0.7395897	0.72330636	0.6539132	957
CG8869	chymotrypsin	0.75348514	0.7245675	0.6523997	831
CG8871	chymotrypsin	0.75066644	0.696529	0.6291706	777
CG7118	serine-type endopeptidase	0.7279424	0.68806565	0.6204272	783
CG6467	serine-type endopeptidase	0.6827621	0.6612262	0.58739495	816
CG5177	trehalase phosphatase	0.69661087	0.6764284	0.6163936	831
PebIII	Ejaculatory bulb protein III	0.6735348	0.6812538	0.6135421	375
Jafrac1	thioredoxin peroxidase 1	0.6901921	0.67821527	0.6128532	585
CG3731	mitochondrial processing peptidase (complex)	0.6781502	0.67086726	0.59851253	1149
CG6295	triacylglycerol lipase	0.698933	0.6635575	0.5943923	1017
CG6783	unknown function	0.7972766	0.7706263	0.72575456	393
CG18594	unknown function	0.7853567	0.7609313	0.71450347	531
CG4800	unknown function	0.76247287	0.73859936	0.682344	519
CG6770	unknown function	0.7347374	0.7237046	0.663708	210
CG7710	unknown function	0.6989024	0.7074764	0.64616084	285
CG16885	unknown function	0.7553575	0.7071962	0.6453053	780
CG11086	unknown function	0.7006948	0.6944669	0.61823046	492
CG13551	unknown function	0.662802	0.6812593	0.62426597	138
CG12726	unknown function	0.6958923	0.7083836	0.6149367	621
CG5468	unknown function	0.73100907	0.66568255	0.6079605	867
CG15704	unknown function	0.65757215	0.6830931	0.6014136	156
CG8756	unknown function	0.66329205	0.66926646	0.59053504	1596
CG1153	unknown function	0.6719674	0.6570362	0.58704245	867
CG2297	unknown function	0.7192255	0.72342896	0.66363025	432
CG7592	unknown function	0.72088313	0.7037263	0.6350077	450

Table 11 (continued, 58 genes out of 128). Reference set of 128 genes for *D. melanogaster* computed by our algorithm. CAI S&al.= computed by Sharp *et al.* ; CAI = computed with this reference set; gCAI = computed with the algorithm in 15 iterations; Len = length of the gene measured in basepairs. Annotation is available at <http://flybase.bio.indiana.edu/> and from the GenBank file.