

Supplemental Table S1. Genome size, protein-coding gene number, total gene number, and gene-coding percentage for the sequenced genomes of eukaryotes, bacteria, archaea, viruses, mitochondria, and chloroplasts estimated based on genome sequences.

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
Eukaryotes				
<i>Anopheles gambiae</i> PEST	230466.9	13260	13749	19.11 (7.0)
<i>Apis mellifera</i> DH4	217194.9	15948	16360	
<i>Arabidopsis thaliana</i> columbia	119707.4	26735	31392	49.99 (28.8)
<i>Aspergillus fumigatus</i> Af293	29384.9	9923	10152	55.39 (50.1)
<i>Aspergillus nidulans</i> FGSC A4	30166.8	9541	9727	59.16
<i>Aspergillus niger</i> CBS 513.88	33975.5	14086	14431	65.46
<i>Aspergillus terreus</i> NIH2624	29164.9	10406	10556	59.92
<i>Bigelowiella natans</i> nucleomorph	372.9	293	340	
<i>Bombyx mori</i>	428700	18510		
<i>Caenorhabditis elegans</i>	100281.2	20056	20935	59.88 (25.8)
<i>Candida albicans</i> SC5314	28504.7	14105	14144	63.97
<i>Candida glabrata</i> CBS 138	12280.4	5181	5397	64.37
<i>Canis familiaris</i>	2445110.1	33202	50110	
<i>Chlamydomonas reinhardtii</i>	121000	15143		
<i>Cryptococcus neoformans</i> neoformans JEC21	19051.9	6475	6609	66.88
<i>Cryptococcus neoformans</i> var. neoformans B-3501A	19699.8	6500	6609	64.93
<i>Cryptosporidium hominis</i> TU502	8171.0	3886	3956	64.91
<i>Cryptosporidium parvum</i> Iowa type II	9098.4	3396	3886	68.90
<i>Cyanidioschyzon merolae</i> 10D	16520.3	5331	5373	(44.90)
<i>Danio rerio</i> Tuebingen	1097524.4	37724	38017	38.23
<i>Debaryomyces hansenii</i> CBS767	12220.8	6893	7108	76.14
<i>Dictyostelium discoideum</i> AX4	33928.5	13322	13362	68.76 (56.3)
<i>Drosophila melanogaster</i>	120401.1	14081	21116	61.37 (15.0)
<i>Encephalitozoon cuniculi</i> GB-M1	2497.5	1996	2053	86.99
<i>Entamoeba histolytica</i> HM-1:IMSS	21581.5	9772	10202	54.53
<i>Eremothecium gossypii</i>	8742.4	4718	4987	80.17 (81.2)
<i>Gallus gallus</i>	1031880.1	15288	15700	(3.1)
<i>Giardia lamblia</i> ATCC 50803	9717.5	6569	6598	94.45
<i>Gibberella zeae</i> PH-1	36487.3	11640	11640	55.64
<i>Guillardia theta</i> nucleomorph	551.3	486	553	90.34
<i>Hemiselmis andersenii</i> CCMP644 nucleomorph	572	472	525	
<i>Homo sapiens</i>	3080436	27727	38612	35.34 (1.20)
<i>Kluyveromyces lactis</i> NRRL Y-1140	10689.2	5331	5502	70.93
<i>Magnaporthe grisea</i> 70-15	41495.9	12832	13146	47.88

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Mus musculus</i> C57BL/6J	2644094.0	39625	60745	35.34 (2.0)
<i>Neurospora crassa</i> OR74A	37097.3	10079	10082	45.49 (37.6)
<i>Oryza sativa</i> (japonica cultivar-group)	370792	26777	28431	
<i>Ostreococcus lucimarinus</i> CCE9901	13204.9	7603	7640	75.31
<i>Ostreococcus tauri</i>	12560	8166		(81.60)
<i>Pan troglodytes</i>	3175580.7	48911	49446	
<i>Paramecium tetraurelia</i> (micronuclear)	72000	39642		(78)
<i>Phaeodactylum tricornutum</i> CCAP1055/1	27400	10402		(52)
<i>Pichia stipitis</i> CBS 6054	15441.2	5816	5816	61.23
<i>Plasmodium falciparum</i> 3D7	22859.5	5268	5298	58.38
<i>Plasmodium yoelii yoelii</i> 17XNL	20087.0	7860	7910	59.12
<i>Rattus norvegicus</i> GK/Ox	2718897.3	27490	38115	31.44 (1.7)
<i>Saccharomyces cerevisiae</i>	12156.6	5869	6276	73.57 (70.5)
<i>Schizosaccharomyces pombe</i> 972h-	12554.3	5093	5364	60.42
<i>Tetrahymena thermophila</i> (macronuclear)	104100	27424	28345	
<i>Thalassiosira pseudonana</i> CCMP1335	34500	11242		
<i>Theileria parva</i> Muguga	8347.6	4079	4159	82.14
<i>Trypanosoma brucei</i> TREU927	20527.4	8772	10253	51.03
<i>Ustilago maydis</i> 521	19638.0	6522	6631	64.33
<i>Volvox carteri</i>	140000	15544		
<i>Yarrowia lipolytica</i> CLIB122	20503.0	6520	7042	47.15
Bacteria				
<i>Acidiphilium cryptum</i> JF-5	3917.1	3551	3552	89.55
<i>Acidothermus cellulolyticus</i> 11B	2445.4	2151	2197	90.39
<i>Acinetobacter baumannii</i> ATCC 17978 (chr)	3976.7	3352	3439	71.00
<i>Acinetobacter</i> sp. ADP1	3598.6	3325	3422	88.85
<i>Actinobacillus pleuropneumoniae</i> 4074	2292.3	2132	2193	86.03
<i>Actinobacillus pleuropneumoniae</i> L20	2274.5	2012	2108	86.00
<i>Actinobacillus succinogenes</i> 130Z	2046.1	1884	1938	89.32
<i>Agrobacterium tumefaciens</i> C58 Cereon	5673.5	5301	5366	89.60
<i>Agrobacterium tumefaciens</i> C58 Dupont	5674.1	5402	5467	88.45
<i>Alkalilimnicola ehrlichei</i> MLHE-1	3272.8	2869	2923	90.66

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Alkaliphilus metalliredigenes</i> QYMF	4410.3	4090	4233	82.44
<i>Anabaena</i> PCC7120	7211.8	6132	6214	82.41
<i>Anabaena variabilis</i> ATCC29413	7105.8	5746	5805	82.51
<i>Anaeromyxobacter dehalogenans</i> 2CP-C	5013.5	4361	4416	91.45
<i>Anaplasma marginale</i> St. Maries	1197.7	949	1005	85.88
<i>Anaplasma phagocytophilum</i> HZ	1471.3	1264	1304	68.62
<i>Aquifex aeolicus</i> VF5	1590.8	1553	1603	92.68
<i>Arthrobacter</i> sp. FB24	5011.6	4524	4591	90.16
<i>Aster yellows</i> witches-broom phytoplasma AYWB	724.0	693	728	74.81
<i>Azoarcus</i> sp. EbN1	4727.3	4598	4668	90.89
<i>Azotobacter vinelandii</i> AvOP	5352.4	4990	5058	88.74
<i>Bacillus anthracis</i> A2012	5093.6	5354	5390	81.08
<i>Bacillus anthracis</i> Ames	5227.3	5311	5439	81.71
<i>Bacillus anthracis</i> Ames 0581	5503.9	5617	5745	81.13
<i>Bacillus anthracis</i> Sterne	5228.7	5287	5415	84.98
<i>Bacillus cereus</i> ATCC 10987	5432.7	5844	6008	85.82
<i>Bacillus cereus</i> ATCC 14579	5427.1	5255	5402	81.91
<i>Bacillus cereus</i> E33L	5300.9	5134	5269	85.14
<i>Bacillus cereus</i> G9241	5934.9	6147	6316	81.24
<i>Bacillus cereus</i> NVH391-98	3915.8	3886	4033	81.99
<i>Bacillus clausii</i> KSM-K16	4303.9	4108	4205	86.95
<i>Bacillus halodurans</i> C-125	4202.4	4066	4172	86.05
<i>Bacillus licheniformis</i> Goettingen	4222.6	4196	4289	87.89
<i>Bacillus licheniformis</i> Novozymes	4222.3	4152	4245	86.78
<i>Bacillus subtilis</i> 168	4214.6	4106	4225	88.30
<i>Bacillus thuringiensis</i> konkukian 97-27	5237.7	5117	5264	85.15
<i>Bacillus weihenstephanensis</i> KBAB4	5602.5	5532	5629	81.97
<i>Bacteroides fragilis</i> NCTC9343	5241.7	4308	4400	89.74
<i>Bacteroides fragilis</i> YCH46	5311.0	4625	4717	90.52
<i>Bacteroides thetaiotaomicron</i> VPI-5482	6293.4	4816	4902	89.88
<i>Bartonella henselae</i> Houston-1	1931.0	1612	1664	76.73
<i>Bartonella quintana</i> Toulouse	1581.4	1308	1357	78.72
<i>Bdellovibrio bacteriovorus</i> HD100	3783.0	3583	3625	93.14
<i>Bifidobacterium longum</i> DJO10A	2375.3	1956	2020	83.21
<i>Bifidobacterium longum</i> NCC2705	2260.3	1729	1802	86.43
<i>Blochmannia floridanus</i>	705.6	589	635	84.91
<i>Blochmannia pennsylvanicus</i> BPEN	791.7	610	692	77.62

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Bordetella bronchiseptica</i> RB50	5339.2	5006	5070	92.34
<i>Bordetella parapertussis</i> 12822	4773.6	4402	4465	91.48
<i>Bordetella pertussis</i> Tohama I	4086.2	3806	3866	90.53
<i>Borrelia burgdorferi</i> B31	1519.9	1639	1677	86.04
<i>Borrelia garinii</i> Pbi	986.9	932	972	92.29
<i>Bradyrhizobium japonicum</i> USDA 110	9105.8	8317	8371	86.76
<i>Bradyrhizobium</i> sp. BTAi1	8422.4	7462	7518	86.45
<i>Brevibacterium linens</i> BL2	4510.7	3982	4038	86.42
<i>Brucella abortus</i> 2308	3278.3	3350	3414	87.68
<i>Brucella abortus</i> 9-941	3286.4	3085	3149	81.82
<i>Brucella melitensis</i> 16M	3294.9	3198	3264	86.67
<i>Brucella suis</i> 1330	3315.2	3273	3337	84.98
<i>Buchnera aphidicola</i> APS	655.7	574	609	87.40
<i>Buchnera aphidicola</i> Bp	618.4	507	542	82.28
<i>Buchnera aphidicola</i> Sg	654.2	555	590	84.14
<i>Burkholderia ambifaria</i> AMMD	7503.6	6539	6617	88.28
<i>Burkholderia cenocepacia</i> AU 1054	7249.5	6475	6555	88.10
<i>Burkholderia cenocepacia</i> HI2424	8139.1	7065	7146	84.74
<i>Burkholderia mallei</i> 10229	6065.2	5133	5201	79.68
<i>Burkholderia mallei</i> 10399	5933.9	4886	4947	80.60
<i>Burkholderia mallei</i> ATCC 23344	5835.5	4764	4831	77.88
<i>Burkholderia mallei</i> FMH	2273.4	2667	2687	78.90
<i>Burkholderia mallei</i> GB8 horse 4	5804.9	4822	4884	80.48
<i>Burkholderia mallei</i> JHU	2190.5	2581	2608	79.81
<i>Burkholderia mallei</i> NCTC 10247	5976.7	5096	5146	79.81
<i>Burkholderia mallei</i> NCTC 10247 (chr1+2)	5848.4	5852	5965	86.40
<i>Burkholderia mallei</i> SAVP1	5283.5	4364	4438	80.98
<i>Burkholderia pseudomallei</i> 1106a (chr1+2)	7089.2	7183	7263	87.44
<i>Burkholderia pseudomallei</i> 1655	7029.5	5410	5465	77.42
<i>Burkholderia pseudomallei</i> 1710a	7319.5	5502	5565	76.60
<i>Burkholderia pseudomallei</i> 1710b	7390.5	5479	5553	76.79
<i>Burkholderia pseudomallei</i> 668	7071.9	5433	5495	78.59
<i>Burkholderia pseudomallei</i> 668 (chr1+2)	7040.4	7230	7307	87.89
<i>Burkholderia pseudomallei</i> K96243	7247.5	5855	5928	83.62
<i>Burkholderia pseudomallei</i> Pasteur	7346.6	5595	5653	77.26
<i>Burkholderia pseudomallei</i> S13	7388.3	5715	5771	78.01
<i>Burkholderia</i> sp. 383	8676.3	7739	7824	87.74

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Burkholderia thailandensis</i> E264	6724.0	5634	5704	86.95
<i>Burkholderia vietnamiensis</i> G4	8411.0	7920	7992	87.75
<i>Burkholderia xenovorans</i> LB400	9703.7	8930	9012	87.57
<i>Burkholderia xenovorans</i> LB400	9731.1	8957	9039	87.52
<i>Caldicellulosiruptor saccharolyticus</i> DSM 8903	2788.3	2602	2647	86.67
<i>Campylobacter coli</i> RM2228	1860.7	1967	2019	92.25
<i>Campylobacter jejuni</i> NCTC 11168	1641.5	1654	1707	95.41
<i>Campylobacter jejuni</i> RM1221	1777.8	1838	1891	91.83
<i>Campylobacter lari</i> RM2100	1562.9	1599	1651	94.54
<i>Campylobacter upsaliensis</i> RM3195	1773.8	1934	1984	93.46
<i>Carboxydotherrmus hydrogenoformans</i> Z-2901	2401.5	2620	2682	91.00
<i>Caulobacter crescentus</i> CB15	4016.9	3737	3794	90.56
<i>Chlamydia muridarum</i> Nigg	1080.5	911	954	90.58
<i>Chlamydia trachomatis</i> A/HAR-13	1052.0	919	960	90.96
<i>Chlamydia trachomatis</i> D/UW-3/CX	1042.5	894	937	90.83
<i>Chlamydomphila abortus</i> S26/3	1144.4	961	1002	90.98
<i>Chlamydomphila caviae</i> GPIC	1181.4	1005	1047	89.78
<i>Chlamydomphila felis</i> Fe/C-56	1173.8	1013	1054	91.53
<i>Chlamydomphila pneumoniae</i> AR39	1229.9	1110	1151	89.22
<i>Chlamydomphila pneumoniae</i> CWL029	1230.2	1052	1118	88.76
<i>Chlamydomphila pneumoniae</i> J138	1226.6	1069	1134	90.05
<i>Chlamydomphila pneumoniae</i> TW-183	1225.9	1113	1154	90.55
<i>Chlorobium limicola</i> DSMZ 245	2761.9	2436	2483	85.87
<i>Chlorobium phaeobacteroides</i> BS1	4444.2	3791	3845	78.11
<i>Chlorobium phaeobacteroides</i> DSMZ 266	3114.3	2792	2840	85.18
<i>Chlorobium tepidum</i> TLS	2154.9	2252	2308	88.46
<i>Chlorobium vibrioforme</i> f. thiosulfatophilum DSMZ 265	1980.2	1747	1798	90.47
<i>Chlorochromatium aggregatum</i>	2572.1	2033	2081	89.07
<i>Chloroflexus aurantiacus</i> J-10-fl	5193.8	3938	3994	82.25
<i>Chromobacterium violaceum</i> ATCC 12472	4751.1	4407	4530	89.40
<i>Chromohalobacter salexigens</i> DSM 3043	3696.6	3319	3404	90.64
<i>Chromohalobacter salexigens</i> DSM 3043	3671.0	3347	3422	90.92
<i>Clostridium acetobutylicum</i> ATCC 824	4132.9	3848	3955	86.93

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Clostridium beijerincki</i> NCIMB 8052	5952.5	5068	5210	80.61
<i>Clostridium difficile</i> 630	4290.3	3742	3971	82.00
<i>Clostridium perfringens</i> 13	3085.7	2723	2849	85.10
<i>Clostridium tetani</i> E88	2873.3	2432	2504	85.96
<i>Clostridium thermocellum</i> ATCC 27405	3895.0	3163	3235	79.84
<i>Clostridium thermocellum</i> ATCC 27405	3843.3	3191	3307	83.00
<i>Colwellia psychrerythraea</i> 34H	5373.2	4910	5114	85.36
<i>Corynebacterium diphtheriae</i> NCTC 13129	2488.6	2320	2389	90.30
<i>Corynebacterium efficiens</i> YS-314	3219.5	2998	3069	90.83
<i>Corynebacterium glutamicum</i> ATCC 13032 Kalinowski	3282.7	3058	3136	88.37
<i>Corynebacterium glutamicum</i> ATCC 13032 Nakagawa	3309.4	3099	3177	87.83
<i>Corynebacterium jeikeium</i> K411	2476.8	2120	2229	89.94
<i>Coxiella burnetii</i> RSA 493	2032.7	2046	2091	84.29
<i>Croceibacter atlanticus</i> HTCC2559	2954.7	2719	2756	92.63
<i>Crocospaera watsonii</i> WH8501	6285.4	5967	6011	78.22
<i>Cupriavidus metallidurans</i> CH34	6887.6	6312	6377	88.89
<i>Cupriavidus necator</i> (Ralstonia eutropha) JMP134	7255.3	6529	6613	88.53
<i>Cytophaga hutchinsonii</i> ATCC 33406	4433.2	3592	3640	89.65
<i>Dechloromonas aromatica</i> RCB	4501.1	4204	4280	92.39
<i>Dehalococcoides ethenogenes</i> 195	1469.7	1580	1629	90.20
<i>Dehalococcoides</i> sp. BAV1	1299.5	1314	1354	89.44
<i>Dehalococcoides</i> sp. CBDB1	1395.5	1458	1508	90.47
<i>Deinococcus geothermalis</i> DSM11300	3164.1	2993	3058	89.53
<i>Deinococcus radiodurans</i> R1	3284.2	3102	3159	87.71
<i>Desulfitobacterium hafniense</i> DCB-2	6083.8	4391	4463	69.23
<i>Desulfitobacterium hafniense</i> Y51	5727.5	5060	5137	85.21
<i>Desulfotalea psychrophila</i> LSv54	3659.6	3236	3343	86.02
<i>Desulfotomaculum reducens</i> MI-1	3467.3	3220	3272	83.57
<i>Desulfovibrio desulfuricans</i> G20	3730.2	3784	3862	91.13
<i>Desulfovibrio vulgaris</i> Hildenborough	3773.2	3531	3615	86.57
<i>Desulfuromonas acetoxidans</i> DSM 684	3529.0	3032	3065	87.31
<i>Ehrlichia canis</i> Jake	1315.0	944	983	72.91
<i>Ehrlichia chaffeensis</i> Arkansas	1176.2	1105	1145	79.94

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Ehrlichia chaffeensis</i> sapulpa	1005.8	805	834	72.99
<i>Ehrlichia ruminantium</i> Gardel	1499.9	980	1019	64.57
<i>Ehrlichia ruminantium</i> str. Welgevonden	1516.0	888	961	62.00
<i>Ehrlichia ruminantium</i> Welgevonden ARC-OVI	1516.4	943	982	63.32
<i>Ehrlichia ruminantium</i> Welgevonden CIRAD	1513.0	976	1015	63.94
<i>Enterococcus faecalis</i> V583	3360.0	3264	3344	86.53
<i>Enterococcus faecium</i> DO	2848.4	2722	2794	84.40
<i>Erwinia carotovora</i> subsp. atroseptica SCRI1043	5064.0	4492	4590	87.15
<i>Erythrobacter litoralis</i> HTCC2594	3052.8	3042	3088	92.05
<i>Escherichia coli</i> K12	4738.8	4359	4468	88.53
<i>Escherichia coli</i> O157:H7 EDL933	5620.5	5449	5571	87.81
<i>Escherichia coli</i> O6 CFT073	5231.4	5379	5489	88.66
<i>Escherichia coli</i> Sakai O157:H7	5591.2	5444	5571	88.19
<i>Escherichia coli</i> UTI89	5180.0	5211	5321	89.64
<i>Exiguobacterium</i> sp. 255-15	2894.1	2976	3073	89.35
<i>Flavobacterium johnsoniae</i> UW101	6069.8	4985	5037	87.26
<i>Francisella tularensis</i> SCHU S4	1892.8	1804	1863	88.25
<i>Frankia</i> sp. CcI3	5381.4	4561	4611	85.50
<i>Frankia</i> sp. CcI3	5433.6	4548	4600	85.56
<i>Frankia</i> sp. EAN1pec	9081.4	7981	8039	85.68
<i>Fusobacterium nucleatum</i> nucleatum ATCC 25586	2174.5	2067	2129	90.05
<i>Fusobacterium nucleatum</i> subsp. vincentii ATCC 49256	2118.3	2250	2327	86.04
<i>Geobacillus kaustophilus</i> HTA426	3592.7	3540	3654	86.10
<i>Geobacter metallireducens</i> GS-15	4011.2	3576	3631	91.55
<i>Geobacter metallireducens</i> GS-15	4011.2	3576	3631	91.55
<i>Geobacter sulfurreducens</i> PCA	3814.1	3447	3502	89.92
<i>Geobacter uraniumreducens</i> Rf4	4882.3	4187	4188	85.94
<i>Gloeobacter violaceus</i> PCC 7421	4659.0	4430	4478	89.31
<i>Gluconobacter oxydans</i> 621H	2922.4	2664	2732	90.02
<i>Haemophilus ducreyi</i> 35000HP	1699.0	1717	1784	86.62
<i>Haemophilus influenzae</i> 86-028NP	1913.4	1792	1869	88.53
<i>Haemophilus influenzae</i> R2846	1824.2	1691	1759	89.11
<i>Haemophilus influenzae</i> R2866	1933.3	1817	1892	89.24
<i>Haemophilus influenzae</i> Rd KW20	1830.1	1709	1785	87.24
<i>Haemophilus somnus</i> 129PT	2008.4	1779	1845	89.59

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Haemophilus somnus</i> 2336	2247.7	2081	2131	87.99
<i>Hahella chejuensis</i> KCTC 2396	7215.3	6778	6860	88.20
<i>Halorhodospira halophila</i> SL1	2674.3	2423	2423	92.10
<i>Haloferox</i> <i>oreni</i> H 168	2464.0	2273	2274	87.08
<i>Helicobacter hepaticus</i> ATCC 51449	1799.1	1875	1939	93.44
<i>Helicobacter pylori</i> 26695	1667.9	1566	1609	90.04
<i>Helicobacter pylori</i> J99	1643.8	1491	1531	90.69
<i>Herminiimonas arsenicoxydans</i>	3424.3	3325	3391	87.00
<i>Idiomarina loihiensis</i> L2TR	2839.3	2628	2697	92.80
<i>Jannaschia</i> sp. CCS1	4404.0	4283	4328	90.77
<i>Jannaschia</i> sp. CCS1	4399.7	4292	4337	90.87
<i>Kineococcus radiotolerans</i> SRS30216	4894.0	4561	4618	90.69
<i>Lactobacillus acidophilus</i> NCFM	1993.6	1864	1944	89.27
<i>Lactobacillus brevis</i> ATCC367	1880.8	1550	1600	72.96
<i>Lactobacillus casei</i> ATCC334	2760.7	2478	2529	71.74
<i>Lactobacillus delbrueckii</i> bulgaricus ATCCBAA-365	1629.4	1543	1622	73.33
<i>Lactobacillus gasseri</i> ATCC33323	1950.2	1695	1757	84.09
<i>Lactobacillus johnsonii</i> NCC 533	1992.7	1820	1915	90.85
<i>Lactobacillus plantarum</i> WCFS1	3308.3	3051	3137	85.25
<i>Lactobacillus reuteri</i> 100-23	2174.3	1972	2037	85.95
<i>Lactobacillus reuteri</i> JCM 1112	1788.9	1707	1758	85.97
<i>Lactobacillus reuteri</i> JCM 1112	1788.9	1707	1758	85.97
<i>Lactobacillus sakei</i> sakei 23K	1884.7	1884	1968	88.73
<i>Lactobacillus salivarius</i> UCC118	2134.0	2017	2116	84.39
<i>Lactococcus lactis</i> cremoris SK11	2613.2	2628	2690	73.03
<i>Lactococcus lactis</i> lactis II1403	2365.6	2266	2346	85.88
<i>Legionella pneumophila</i> Lens	3345.7	2947	2999	88.58
<i>Legionella pneumophila</i> Paris	3503.6	3082	3134	88.41
<i>Legionella pneumophila</i> Philadelphia 1	3397.8	2942	2997	88.67
<i>Leifsonia xyli</i> xyli CTCB07	2584.2	2030	2078	70.42
<i>Leptospira interrogans</i> serovar Copenhageni Fiocruz	4627.4	3658	3700	74.98
<i>Leptospira interrogans</i> serovar Lai 56601	4691.2	4725	4766	78.18
<i>Leuconostoc mesenteroides</i> ATCC 8293	1976.6	1869	1921	85.68
<i>Listeria innocua</i> Clip11262	3093.1	3061	3145	89.97
<i>Listeria monocytogenes</i> 4b F2365	2905.3	2821	2906	89.43
<i>Listeria monocytogenes</i> EGD-e	2944.5	2855	2940	90.22

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Listeria monocytogenes</i> str. 1/2a F6854	2950.3	2967	3052	89.53
<i>Listeria monocytogenes</i> str. 4b H7858	2972.3	3111	3191	89.65
<i>Lyngbya aestuarii</i>	7037.5	2505	6142	87.36
<i>Magnetococcus</i> sp. MC-1	4628.7	3718	3770	86.10
<i>Magnetospirillum magneticum</i> AMB-1	4967.1	4559	4763	88.36
<i>Magnetospirillum magnetotacticum</i> MS-1	4503.3	4306	4357	89.72
<i>Mannheimia succiniciproducens</i> MBEL55E	2314.1	2384	2463	90.81
<i>Marinobacter aquaeolei</i> VT8	4648.0	4158	4210	89.70
<i>Mesoplasma florum</i> L1	793.2	683	716	94.19
<i>Mesorhizobium loti</i> MAFF303099	7596.3	7281	7341	85.91
<i>Mesorhizobium</i> sp. BNC1	4978.0	4480	4537	87.93
<i>Methylobacillus flagellatus</i> KT	2851.3	2505	2553	87.36
<i>Methylobacillus flagellatus</i> KT	2971.5	2759	2811	90.73
<i>Methylococcus capsulatus</i> Bath	3304.6	2959	3011	89.60
<i>Moorella thermoacetica</i> ATCC 39073	2628.8	2523	2577	87.86
<i>Mycobacterium avium</i> paratuberculosis k10	4829.8	4350	4399	91.51
<i>Mycobacterium bovis</i> AF2122/97	4345.5	3953	4002	91.05
<i>Mycobacterium leprae</i> TN	3268.2	2720	2768	76.87
<i>Mycobacterium</i> sp. JLS	6048.4	5739	5845	92.00
<i>Mycobacterium tuberculosis</i> CDC1551	4403.8	4189	4238	90.37
<i>Mycobacterium tuberculosis</i> H37Rv	4411.5	3999	4047	91.09
<i>Mycoplasma gallisepticum</i> R	996.4	726	765	88.43
<i>Mycoplasma genitalium</i> G-37	580.1	480	519	91.47
<i>Mycoplasma hyopneumoniae</i> 232	892.8	691	724	90.51
<i>Mycoplasma hyopneumoniae</i> 7448	920.1	663	726	86.69
<i>Mycoplasma hyopneumoniae</i> J	897.4	665	728	88.41
<i>Mycoplasma mobile</i> 163K	777.1	635	666	91.64
<i>Mycoplasma mycoides</i> SC PG1	1211.7	1016	1052	82.45
<i>Mycoplasma penetrans</i> HF-2	1358.6	1037	1072	89.09
<i>Mycoplasma pneumoniae</i> M129	816.4	688	728	88.69
<i>Mycoplasma pulmonis</i> UAB CTIP	963.9	782	814	90.73
<i>Mycoplasma synoviae</i> 53	799.5	672	747	90.64
<i>Neisseria gonorrhoeae</i> FA 1090	2153.9	2002	2069	79.15
<i>Neisseria meningitidis</i> MC58	2272.4	2025	2097	78.70
<i>Neisseria meningitidis</i> Z2491	2184.4	2121	2191	83.79

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Neorickettsia sennetsu</i> Miyayama	859.0	932	968	88.02
<i>Nitrobacter hamburgensis</i> X14	5010.3	4669	4728	83.19
<i>Nitrobacter hamburgensis</i> X14	5011.5	4673	4731	83.91
<i>Nitrobacter winogradskyi</i> Nb-255	3402.1	3143	3195	84.90
<i>Nitrosococcus oceani</i> ATCC 19707	3522.1	3132	3183	86.77
<i>Nitrosomonas europaea</i> ATCC 19718	2812.1	2574	2619	88.49
<i>Nitrosomonas eutropha</i> C71	2712.0	2544	2588	86.73
<i>Nitrospira multififormis</i> ATCC25196	3234.3	2827	2873	85.64
<i>Nocardia farcinica</i> IFM 10152	6292.3	5936	5999	90.34
<i>Nocardioides</i> sp. JS614	5394.1	4902	4954	89.35
<i>Nostoc punctiforme</i> PCC73102	9059.2	7364	7465	80.91
<i>Nostoc</i> sp. PCC 7120	7000.0	5366	5430	82.00
<i>Novosphingobium aromaticivorans</i> DSM 12444	4226.6	3765	3823	90.79
<i>Novosphingobium aromaticivorans</i> DSM 12444	3561.6	3338	3404	92.20
<i>Oceanicaulis alexandrii</i> HTCC2633	3168.2	3029	3077	90.53
<i>Oceanobacillus iheyensis</i> HTE831	3630.5	3496	3588	85.38
<i>Oenococcus oeni</i> PSU-1	1782.8	1788	1834	84.14
<i>Onion yellows</i> phytoplasma OY-M	860.6	754	792	74.08
<i>Parachlamydia</i> sp. UWE25	2414.5	2031	2075	82.75
<i>Paracoccus denitrificans</i> PD1222	5175.7	5101	5157	91.75
<i>Pasteurella multocida</i> subsp. multocida Pm70	2257.5	2014	2090	90.34
<i>Pediococcus pentosaceus</i> ATCC25745	1814.6	1660	1722	86.08
<i>Pelagibacter ubique</i> HTCC1062	1308.8	1354	1406	96.06
<i>Pelobacter carbinolicus</i> DSM 2380	3662.3	3145	3205	86.72
<i>Pelobacter propionicus</i> DSM 2379	4466.7	3828	3896	83.47
<i>Pelodictyon luteolum</i> DSMZ 273(T)	2364.8	2133	2187	90.55
<i>Pelodictyon phaeoclathratiforme</i> BU-1	3000.2	2762	2814	86.28
<i>Photobacterium profundum</i> SS9	6403.3	5480	5697	82.65
<i>Photorhabdus luminescens</i> laumondii TTO1	5689.0	4905	5012	84.24
<i>Polaromonas naphthalenivorans</i> CJ2	5340.2	4973	5022	88.90
<i>Polaromonas</i> sp. JS666	5912.3	5662	5709	88.97
<i>Polaromonas</i> sp. JS666	5898.7	5569	5616	88.60
<i>Porphyromonas gingivalis</i> W83	2343.5	1909	1974	83.33
<i>Prochlorococcus marinus</i> MED4	1658.0	1716	1756	88.48
<i>Prochlorococcus marinus</i> MIT 9313	2410.9	2273	2324	82.28

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Prochlorococcus marinus</i> str. MIT 9211	1839.0	2123	2168	89.00
<i>Prochlorococcus marinus</i> str. MIT 9301	1641.9	1907	1963	90.00
<i>Prochlorococcus marinus</i> str. MIT 9303	2682.7	2997	3136	84.00
<i>Prochlorococcus marinus</i> str. MIT 9312	1709.0	1809	1855	89.00
<i>Prochlorococcus marinus</i> str. NATL2A	1843.0	1890	1938	85.00
<i>Prochlorococcus marinus</i> subsp. <i>marinus</i> CCMP1375	1751.1	1882	1926	89.13
<i>Prochlorococcus marinus</i> subsp. <i>marinus</i> str. SS120 (CCMP1375)	1751.0	1882	1930	88.00
<i>Prochlorococcus marinus</i> subsp. <i>pastoris</i> str. CCMP1986(MED4)	1668.0	1712	1760	87.00
<i>Prochlorococcus</i> sp. CC9902 (coastal)	2234.8	2321	2372	90.31
<i>Propionibacterium acnes</i> KPA171202	2560.3	2297	2351	89.87
<i>Prosthecochloris aestuarii</i> SK413	2563.2	2315	2359	87.41
<i>Pseudoalteromonas atlantica</i> T6c	5095.0	4231	4297	86.68
<i>Pseudoalteromonas haloplanktis</i> TAC 125	3850.3	3487	3621	88.39
<i>Pseudomonas aeruginosa</i> PAO1	6264.4	5566	5642	89.59
<i>Pseudomonas aeruginosa</i> UCBPP-PA14	6524.1	5688	5754	85.93
<i>Pseudomonas fluorescens</i> Pf-5	7074.9	6137	6295	89.06
<i>Pseudomonas fluorescens</i> PfO-1	6438.4	5745	5837	90.42
<i>Pseudomonas putida</i> F1	5925.1	5251	5341	89.13
<i>Pseudomonas putida</i> KT2440	6181.9	5350	5445	87.19
<i>Pseudomonas syringae</i> B728a	6093.7	5137	5217	88.50
<i>Pseudomonas syringae</i> phaseolicola 1448A	6112.4	5170	5312	84.15
<i>Pseudomonas syringae</i> pv. <i>tomato</i> DC3000	6538.3	5608	5688	85.38
<i>Psychrobacter arcticum</i> 273-4	2650.7	2147	2208	82.06
<i>Psychrobacter cryohalolentis</i> K5	3081.1	2485	2538	83.08
<i>Psychrobacter cryohalolentis</i> K5	3101.1	2515	2575	83.80
<i>Psychrobacter</i> sp. PRwf-1	2953.1	2350	2394	83.77
<i>Ralstonia solanacearum</i> GMI1000	5810.9	5120	5189	87.78
<i>Rhizobium etli</i> CFN 42	4382.0	4043	4126	91.89

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Rhizobium etli</i> CFN 42	6159.0	5627	5686	86.42
<i>Rhodobacter sphaeroides</i> 2.4.1	4603.1	4304	4367	88.28
<i>Rhodobacter sphaeroides</i> ATCC 17025	4479.4	4248	4296	88.68
<i>Rhodobacter sphaeroides</i> ATCC 17029	4462.0	3971	4019	88.28
<i>Rhodobacter sphaeroides</i> ATCC 17029 (chr1+2)	4366.8	4025	4135	88.16
<i>Rhodoferrax ferrireducens</i> DSM 15236	4926.7	4378	4427	88.56
<i>Rhodoferrax ferrireducens</i> DSM 15236	4969.8	4495	4546	90.32
<i>Rhodopirellula baltica</i> SH 1	7145.6	7321	7403	94.95
<i>Rhodopirellula baltica</i> SH1	7146.0	7325	7404	94.00
<i>Rhodopseudomonas palustris</i> BisA53	5502.4	4913	4969	86.33
<i>Rhodopseudomonas palustris</i> BisB18	5505.0	4949	5007	86.01
<i>Rhodopseudomonas palustris</i> BisB18	5513.8	4943	4999	86.23
<i>Rhodopseudomonas palustris</i> BisB5	4890.4	4385	4443	87.09
<i>Rhodopseudomonas palustris</i> BisB5	4892.7	4418	4475	87.24
<i>Rhodopseudomonas palustris</i> CGA009	5467.6	4840	4897	87.53
<i>Rhodopseudomonas palustris</i> HaA2	5331.5	4683	4736	87.62
<i>Rhodopseudomonas palustris</i> HaA2	5331.7	4712	4765	87.58
<i>Rhodoseudomonas palustris</i> CGA009	5459.0	4813	4891	87.00
<i>Rhodospirillum rubrum</i> ATCC 11170	4406.6	3850	3917	88.82
<i>Rickettsia akari</i> Hartford	1231.1	1217	1253	77.43
<i>Rickettsia bellii</i> RML369-C	1522.1	1429	1466	85.39
<i>Rickettsia conorii</i> Malish 7	1268.8	1374	1410	80.84
<i>Rickettsia felis</i> URRWXCal2	1587.2	1512	1548	83.82
<i>Rickettsia prowazekii</i> Madrid E	1111.5	834	870	76.03
<i>Rickettsia rickettsii</i>	1257.7	1311	1346	78.50
<i>Rickettsia sibirica</i> 246	1250.0	1234	1270	78.13
<i>Rickettsia typhi</i> Wilmington	1111.5	838	874	75.92
<i>Roseovarius nubinhibens</i> ISM	3668.7	3547	3597	89.58
<i>Rubrivivax gelatinosus</i> PM1	4643.7	4475	4548	91.89
<i>Rubrobacter xylanophilus</i> DSM 9941	3299.4	2990	3039	86.62
<i>Saccharophagus degradans</i> 2-40	5057.5	4017	4064	87.03
<i>Saccharophagus degradans</i> 2-40	5057.5	4017	4064	86.99
<i>Saccharopolyspora erythraea</i> NRRL 2338	8212.8	7191	7264	84.00
<i>Salinibacter ruber</i> DSM 13855	3587.3	2833	2880	84.88

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Salmonella enterica</i> choleraesuis SC-B67	4944.0	4666	4773	84.26
<i>Salmonella enterica</i> CT18	4809.0	4395	4711	83.00
<i>Salmonella enterica</i> Paratyphi-A SARB42	4585.2	4093	4197	83.22
<i>Salmonella enterica</i> Typhi CT18	5133.7	4980	5082	88.07
<i>Salmonella enterica</i> Typhi Ty2	4792.0	4323	4423	83.08
<i>Salmonella typhimurium</i> LT2	4951.4	4554	4692	87.28
<i>Shewanella amazonensis</i> SB2B	4264.5	3577	3702	87.83
<i>Shewanella baltica</i> OS155	5084.3	4334	4458	82.56
<i>Shewanella baltica</i> OS155 (chr)	5127.4	4307	4521	83.00
<i>Shewanella denitrificans</i> OS217	4491.5	3722	3840	84.27
<i>Shewanella denitrificans</i> OS217	4545.9	3778	3896	85.68
<i>Shewanella frigidimarina</i> NCMB400	4782.4	4002	4124	84.41
<i>Shewanella loihica</i> PV-4	4602.6	3859	3993	85.00
<i>Shewanella oneidensis</i> MR-1	5131.4	4778	4908	83.89
<i>Shewanella putrefaciens</i> CN-32	4577.7	3981	4095	85.15
<i>Shewanella</i> sp. ANA-3	5100.7	4241	4353	85.99
<i>Shewanella</i> sp. MR-4	4579.2	3852	3928	85.44
<i>Shewanella</i> sp. MR-7	4546.4	3824	3920	85.56
<i>Shewanella</i> sp. PV-4	4474.5	3704	3811	85.72
<i>Shewanella</i> sp. W3-18-1	4754.0	4117	4241	84.35
<i>Shigella boydii</i> sv4 Sb227	4646.5	4290	4403	81.10
<i>Shigella dysenteriae</i> sv1 Sd197	4552.0	4500	4607	77.88
<i>Shigella flexneri</i> 2a 2457T	4599.4	4073	4195	78.05
<i>Shigella flexneri</i> 2a 301	4828.8	4703	4822	86.98
<i>Shigella sonnei</i> Ss046	5039.7	4462	4581	81.02
<i>Silicibacter pomeroyi</i> DSS-3	4601.1	4252	4314	89.92
<i>Silicibacter</i> sp. TM1040	4198.3	3860	3932	88.15
<i>Sinorhizobium meliloti</i> 1021	6691.7	6205	6270	86.25
<i>Sinorhizobium meliloti</i> 1021	6691.7	6205	6269	86.25
<i>Solibacter usitatus</i> Ellin6076	9949.9	8098	8157	90.92
<i>Sphingopyxis alaskensis</i> RB2256	3343.4	3199	3247	91.37
<i>Staphylococcus aureus</i> bovine RF122	2742.5	2589	2665	84.81
<i>Staphylococcus aureus</i> COL	2813.9	2618	2690	83.69
<i>Staphylococcus aureus</i> JH1	2879.6	2681	2737	83.69
<i>Staphylococcus aureus</i> JH9	2862.9	2673	2674	83.53
<i>Staphylococcus aureus</i> MRSA252	2902.6	2744	2820	84.79
<i>Staphylococcus aureus</i> MSSA476	2820.5	2643	2722	84.81
<i>Staphylococcus aureus</i> Mu50	2903.6	2733	2809	84.59
<i>Staphylococcus aureus</i> MW2	2841.1	2659	2738	84.60

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Staphylococcus aureus</i> N315	2839.5	2624	2702	84.36
<i>Staphylococcus aureus</i> NCTC 832	2821.4	2892	2969	85.71
<i>Staphylococcus aureus</i> USA300	2917.5	2604	2673	82.94
<i>Staphylococcus epidermidis</i> ATCC 12228	2564.6	2485	2561	84.13
<i>Staphylococcus epidermidis</i> RP62A	2643.8	2526	2606	83.70
<i>Staphylococcus haemolyticus</i> JCSC1435	2685.0	2678	2813	86.93
<i>Staphylococcus saprophyticus</i> saprophyticus ATCC 15305	2577.9	2514	2595	84.70
<i>Streptococcus agalactiae</i> 2603V/R	2160.3	2124	2225	88.14
<i>Streptococcus agalactiae</i> Ia A909	2127.8	1996	2097	87.85
<i>Streptococcus agalactiae</i> NEM316	2211.5	2134	2235	90.17
<i>Streptococcus mutans</i> UA159	2030.9	1960	2040	87.06
<i>Streptococcus pneumoniae</i> R6	2038.6	2043	2113	87.57
<i>Streptococcus pneumoniae</i> TIGR4	2160.8	2094	2164	83.54
<i>Streptococcus pyogenes</i> M1 GAS	1852.4	1696	1774	85.34
<i>Streptococcus pyogenes</i> M1 MGAS5005	1838.6	1865	2017	88.22
<i>Streptococcus pyogenes</i> M28 MGAS6180	1897.6	1894	2044	88.48
<i>Streptococcus pyogenes</i> MGAS10394	1899.9	1886	1973	88.65
<i>Streptococcus pyogenes</i> MGAS315	1900.5	1865	1946	87.23
<i>Streptococcus pyogenes</i> MGAS8232	1895.0	1845	1930	86.70
<i>Streptococcus pyogenes</i> SSI-1	1894.3	1861	1930	86.13
<i>Streptococcus suis</i> 89/1591	1978.2	1896	1938	86.10
<i>Streptococcus thermophilus</i> CNRZ 1066	1796.2	1915	2000	85.47
<i>Streptococcus thermophilus</i> LMD-9	1842.1	1850	1919	76.15
<i>Streptococcus thermophilus</i> LMG 18311	1796.8	1889	1974	85.52
<i>Streptomyces avermitilis</i> MA-4680	9119.9	7673	7761	86.39
<i>Streptomyces coelicolor</i> A3(2)	9054.8	8215	8299	89.13
<i>Sulfitobacter</i> sp. EE-36	3547.2	3474	3529	90.89
<i>Sulfitobacter</i> sp. NAS-14.1	4002.1	3962	4014	89.92
<i>Symbiobacterium thermophilum</i> IAM 14863	3566.1	3337	3454	88.01
<i>Synechococcus elongatus</i> BP-1	2593.9	2475	2520	89.80
<i>Synechococcus elongatus</i> PCC7942	2695.0	2611	2664	88.00
<i>Synechococcus</i> sp. CC9605	2511.0	2638	2749	86.00
<i>Synechococcus</i> sp. CC9902	2235.0	2304	2355	89.00

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Synechococcus</i> sp. JA-2-3B'a(2-13)	3047.0	2862	2942	85.00
<i>Synechococcus</i> sp. JA-3-3Ab	2933.0	2760	2897	84.00
<i>Synechococcus</i> sp. PCC 6301	2696.3	2525	2576	87.98
<i>Synechococcus</i> sp. RS9917	2580.0	2770	2820	92.00
<i>Synechococcus</i> sp. WH 5701	3043.0	3346	3401	87.00
<i>Synechococcus</i> sp. WH 7805	2620.0	2883	2394	90.00
<i>Synechococcus</i> sp. WH8102	2434.4	2526	2576	90.45
<i>Synechocystis</i> sp. PCC 6803	3947.0	3564	3614	86.69
<i>Syntrophobacter fumaroxidans</i> MPOB	4848.8	4031	4084	82.96
<i>Syntrophomonas wolfei</i> Goettingen	2845.8	2466	2524	83.33
<i>Syntrophus aciditrophicus</i>	3179.3	3168	3219	88.68
<i>Thermoanaerobacter ethanolicus</i> 39E	2282.7	2254	2311	87.84
<i>Thermoanaerobacter tengcongensis</i> MB4	2689.4	2588	2652	87.42
<i>Thermobifida fusca</i> YX	3642.2	3117	3182	85.74
<i>Thermotoga maritima</i> MSB8	1860.7	1846	1895	93.88
<i>Thermus thermophilus</i> HB27	2127.5	2210	2263	94.77
<i>Thermus thermophilus</i> HB8	1859.1	1988	2042	95.70
<i>Thiobacillus denitrificans</i> sp. ATCC 25259	2909.8	2827	2876	92.78
<i>Thiomicrospira crunogena</i> XCL-2	2427.7	2201	2254	90.63
<i>Thiomicrospira denitrificans</i> ATCC 33889	2201.6	2104	2160	93.80
<i>Treponema denticola</i> ATCC 35405	2843.2	2767	2817	91.93
<i>Treponema pallidum</i> Nichols	1138.0	1031	1082	93.15
<i>Trichodesmium erythraeum</i> IMS101	8191.4	4759	4801	58.12
<i>Tropheryma whipplei</i> TW08/27	925.9	784	838	85.61
<i>Tropheryma whipplei</i> Twist	927.3	808	861	86.34
<i>Ureaplasma parvum</i> serovar 3 ATCC 700970	751.7	611	647	92.43
<i>Vibrio cholerae</i> O1 biovar eltor N16961	4033.5	3828	3951	87.42
<i>Vibrio fischeri</i> ES114	4238.2	3747	3903	87.95
<i>Vibrio parahaemolyticus</i> RIMD 2210633	5165.8	4832	4992	87.86
<i>Vibrio vulnificus</i> CMCP6	5126.8	4537	4676	85.73
<i>Vibrio vulnificus</i> YJ016	5260.1	5028	5168	89.25
<i>Wigglesworthia glossinidia</i> endosymbiont of <i>Glossina</i>	697.7	611	651	88.47
<i>Wolbachia endosymbiont</i> of <i>Drosophila ananassae</i>	1440.8	1802	1833	75.78

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Wolbachia endosymbiont</i> of <i>Drosophila simulans</i>	1063.1	760	782	65.02
<i>Wolbachia endosymbiont</i> strain TRS of <i>Brugia malayi</i>	1080.1	805	842	67.60
<i>Wolbachia pipientis</i> wMel	1267.8	1195	1232	80.59
<i>Wolinella succinogenes</i> DSM 1740	2110.4	2044	2090	94.46
<i>Xanthobacter autotrophicus</i> Py2	5580.9	5041	5042	88.19
<i>Xanthomonas axonopodis</i> citri 306	5274.2	4427	4487	86.01
<i>Xanthomonas campestris</i> campestris 8004	5148.7	4273	4389	84.89
<i>Xanthomonas campestris</i> campestris ATCC 33913	5076.2	4181	4241	84.80
<i>Xanthomonas campestris</i> vesicatoria 85-10	5420.2	4726	4786	87.35
<i>Xanthomonas oryzae</i> KACC10331	4941.4	4637	4697	85.98
<i>Xanthomonas oryzae</i> MAFF 311018	4940.2	4372	4431	84.15
<i>Xylella fastidiosa</i> 9a5c	2731.8	2832	2888	83.64
<i>Xylella fastidiosa</i> Ann-1	2622.3	2359	2411	80.73
<i>Xylella fastidiosa</i> Dixon	5118.9	4662	4757	79.14
<i>Xylella fastidiosa</i> Temecula1	2521.1	2036	2091	78.32
<i>Yersinia pestis</i> biovar Medievalis 91001	4803.2	4142	4237	82.08
<i>Yersinia pestis</i> CO92	4829.9	4217	4307	84.23
<i>Yersinia pestis</i> KIM	4600.8	4090	4187	83.19
<i>Yersinia pseudotuberculosis</i> IP 32953	4840.9	4116	4226	84.27
<i>Zymomonas mobilis</i> ZM4	2056.4	1998	2086	86.72
Archaea				
<i>Aeropyrum pernix</i> K1	1669.7	2694	2751	89.17
<i>Aeropyrum pernix</i> K1	1669.7	2763	2815	90.44
<i>Archaeoglobus fulgidus</i> DSM 4304	2178.4	2407	2456	91.20
<i>Archaeoglobus fulgidus</i> VC-16	2178.4	2466	2515	93.62
<i>Cenarchaeum symbiosum</i>	2023.7	2011	2059	91.52
<i>Ferroplasma acidarmanus</i> fer1	1971.4	1695	1742	80.26
<i>Haloarcula marismortui</i> ATCC 43049	4274.6	4240	4300	84.96
<i>Haloarcula marismortui</i> ATCC 43049	4274.6	4322	4382	86.39
<i>Halobacterium salinarium</i> NRC-1	2571.0	2675	2726	87.41
<i>Halobacterium salinarum</i> NRC-1	2571.0	2605	2657	85.20
<i>Methanobacterium thermoautotrophicum</i> Delta H	1751.4	1846	1891	91.12

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Methanocaldococcus jannaschii</i> DSM 2661	1739.9	1770	1813	87.23
<i>Methanocaldococcus jannaschii</i> DSM 2661	1739.9	1789	1832	88.75
<i>Methanococcoides burtonii</i> DSM 6242	2575.0	2434	2495	85.69
<i>Methanococcoides burtonii</i> DSM6242	2597.4	2490	2551	84.21
<i>Methanococcus maripaludis</i> C5	1780.8	1813	1880	86.00
<i>Methanococcus maripaludis</i> S2	1661.1	1722	1770	89.87
<i>Methanococcus maripaludis</i> S2	1661.1	1732	1780	90.03
<i>Methanoculleus marisnigri</i>	2460.2	2418	2466	87.51
<i>Methanoculleus marisnigri</i> JR1	2478.1	2489	2555	87.00
<i>Methanopyrus kandleri</i> AV19	1695.0	1691	1734	89.09
<i>Methanopyrus kandleri</i> AV19	1695.0	1727	1765	90.24
<i>Methanosarcina acetivorans</i> C2A	5751.5	4540	4609	74.28
<i>Methanosarcina acetivorans</i> C2A	5751.5	4721	4731	76.77
<i>Methanosarcina barkeri</i> Fusaro	4873.8	3759	3830	71.46
<i>Methanosarcina mazei</i> Go1	4096.3	3371	3438	75.83
<i>Methanosarcina mazei</i> Go1	4096.3	3398	3465	76.51
<i>Methanosphaera stadtmanae</i> DSM 3091	1767.4	1534	1586	85.20
<i>Methanospirillum hungatei</i> JF-1	3535.0	3307	3369	87.10
<i>Methanospirillum hungatei</i> JF-1	3544.7	3238	3303	88.62
<i>Methanothermobacter thermautotrophicus</i> Delta H	1751.4	1869	1915	90.91
<i>Methaosaeta thermophila</i> PT	1861.6	1718	1761	84.05
<i>Nanoarchaeum equitans</i> Kin4-M	490.9	536	579	93.39
<i>Nanoarchaeum equitans</i> Kin4-M	490.9	556	573	93.07
<i>Natronomonas pharaonis</i> DSM 2160	2749.7	2843	2892	90.61
<i>Picrophilus torridus</i> DSM 9790	1545.9	1535	1628	91.92
<i>Picrophilus torridus</i> DSM 9790	1545.9	1551	1598	92.41
<i>Pyrobaculum aerophilum</i> IM2	2222.4	2605	2655	88.78
<i>Pyrobaculum aerophilum</i> IM2	2222.4	2575	2614	89.24
<i>Pyrobaculum calidifontis</i> JCM 11548	2009.3	2149	2200	90.00
<i>Pyrococcus abyssi</i> GE5	1765.1	1784	1881	92.04
<i>Pyrococcus abyssi</i> GE5	1768.6	1904	1955	93.78
<i>Pyrococcus furiosus</i> DSM 3638	1908.3	2065	2115	91.38
<i>Pyrococcus furiosus</i> DSM 3638	1908.3	2069	2171	92.66
<i>Pyrococcus horikoshii</i> OT3	1738.5	2061	2112	91.07
<i>Pyrococcus horikoshii</i> OT3	1738.5	2196	2297	93.31

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Staphylothermus marinus</i> F1	1570.5	1570	1646	86.00
<i>Sulfolobus acidocaldarius</i> DSM 639	2226.0	2223	2322	85.31
<i>Sulfolobus acidocaldarius</i> DSM 639	2226.0	2285	2336	87.67
<i>Sulfolobus solfataricus</i> P2	2992.2	2994	3050	84.32
<i>Sulfolobus solfataricus</i> P2	2992.2	3048	3104	86.43
<i>Sulfolobus tokodaii</i> 7	2694.8	2826	2878	83.36
<i>Sulfolobus tokodaii</i> str. 7	2694.8	2915	2964	85.52
<i>Thermococcus kodakaraensis</i> KOD1	2088.7	2306	2358	92.05
<i>Thermococcus kodakaraensis</i> KOD1	2088.7	2297	2349	92.09
<i>Thermofilum pendens</i>	1811.4	1892	1932	90.99
<i>Thermoplasma acidophilum</i> DSM 1728	1564.9	1478	1549	87.02
<i>Thermoplasma acidophilum</i> DSM 1728	1564.9	1527	1575	88.51
<i>Thermoplasma volcanium</i> GSS1	1584.8	1526	1575	86.19
<i>Thermoplasma volcanium</i> GSS1	1584.8	1561	1610	88.26
Virus				
Acholeplasma phage L2	12.0	14	14	81.31
Acholeplasma phage MV-L1	4.5	4	4	46.60
Actinoplanes phage phiAsp2	58.6	76	76	92.23
Aeromonas phage 31	173.0	247	263	92.84
Bacillus clarkii bacteriophage BCJA1c	41.1	58	58	92.66
Bacillus phage GA-1	21.1	35	35	92.46
Bacillus phage phi29	19.4	27	27	93.88
Bacillus thuringiensis bacteriophage Bam35c	14.9	32	32	93.49
Bacillus thuringiensis phage GIL16c	14.8	31	31	94.09
Bacteriophage 11b	36.1	65	65	90.15
Bacteriophage 187	39.6	77	77	93.55
Bacteriophage 2638A	41.3	57	57	92.68
Bacteriophage 29	42.8	75	75	92.15
Bacteriophage 37	43.7	77	77	95.58
Bacteriophage 3A	43.1	67	67	93.95
Bacteriophage 42e	45.9	79	79	93.38
Bacteriophage 44RR2.8t	173.6	252	269	93.28
Bacteriophage 47	44.8	72	72	93.99
Bacteriophage 52A	41.7	65	65	91.64
Bacteriophage 53	43.9	79	79	93.42
Bacteriophage 55	41.9	77	77	93.37

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
Bacteriophage 66	18.2	27	27	91.85
Bacteriophage 69	42.7	76	76	93.02
Bacteriophage 71	43.1	72	72	94.55
Bacteriophage 77	41.7	69	69	92.12
Bacteriophage 85	44.3	78	78	91.80
Bacteriophage 88	43.2	72	72	94.35
Bacteriophage 92	42.4	74	74	95.17
Bacteriophage 933W	61.7	80	83	88.07
Bacteriophage 96	43.6	79	79	91.98
Bacteriophage A118	40.8	72	72	94.19
Bacteriophage Aaphi23	43.0	66	66	92.42
Bacteriophage Aeh1	233.2	352	377	94.04
Bacteriophage AP205	4.3	4	4	90.96
Bacteriophage B103	18.6	17	17	84.76
Bacteriophage B3	38.4	59	59	97.13
Bacteriophage bIL170	31.8	64	65	91.44
Bacteriophage bIL285	35.5	62	62	93.03
Bacteriophage bIL286	41.8	61	61	93.09
Bacteriophage bIL309	36.9	56	56	89.47
Bacteriophage bIL310	15.0	29	29	77.29
Bacteriophage bIL311	14.5	22	23	61.98
Bacteriophage bIL312	15.2	27	28	75.55
Bacteriophage D3112	37.6	55	55	93.84
Bacteriophage EJ-1	42.9	73	73	92.25
Bacteriophage EW	45.3	77	77	92.04
Bacteriophage Felix 01	86.2	247	269	92.91
Bacteriophage G1	138.7	214	218	88.69
Bacteriophage HK620	38.3	58	58	88.62
Bacteriophage IN93	19.6	35	35	83.53
Bacteriophage K139	33.1	44	44	92.10
Bacteriophage KS7	40.8	56	56	74.94
Bacteriophage KVP40	244.8	381	410	91.06
Bacteriophage L-413C	30.7	40	40	93.17
Bacteriophage L5	2.4	8	8	78.97
Bacteriophage lambda	48.5	71	71	87.83
Bacteriophage Mx8	49.5	86	86	93.40
Bacteriophage P27	42.6	58	60	89.84
Bacteriophage P4	11.6	14	14	84.25
Bacteriophage phBC6A51	61.4	75	75	82.74
Bacteriophage phBC6A52	38.5	49	49	80.07
Bacteriophage phi AT3	39.2	55	55	89.65

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
Bacteriophage phi CTX	35.6	47	47	91.27
Bacteriophage phi ETA	43.1	66	66	92.79
Bacteriophage phi JL001	63.6	90	90	95.32
Bacteriophage phi LC3	32.2	51	51	92.46
Bacteriophage phi1026b	54.9	83	83	94.00
Bacteriophage phi-105	39.3	51	51	89.23
Bacteriophage phi-12	13.2	15	15	85.87
Bacteriophage phi3626	33.5	50	50	92.54
Bacteriophage phi-8	15.0	16	16	89.33
Bacteriophage phi-BT1	41.8	55	58	88.36
Bacteriophage phi-C31	41.5	53	55	90.01
Bacteriophage phiE125	53.4	71	71	89.70
Bacteriophage phig1e	42.3	50	52	76.25
Bacteriophage phiKMV	42.5	48	48	90.81
Bacteriophage phiKO2	51.6	64	64	92.50
Bacteriophage phiMFV1	15.1	17	17	91.61
Bacteriophage phiYeO3-12	39.6	59	59	91.36
Bacteriophage PSP3	30.6	42	42	91.69
Bacteriophage PT1028	15.6	22	22	77.42
Bacteriophage PY54	46.3	67	67	89.84
Bacteriophage r1t	33.3	50	50	91.26
Bacteriophage RM 378	129.9	146	147	93.44
Bacteriophage ROSA	43.2	74	74	92.86
Bacteriophage sk1	28.5	56	57	89.53
Bacteriophage SPBc2	134.4	185	185	86.22
Bacteriophage S-PM2	196.3	239	264	94.01
Bacteriophage SPP1	44.0	106	106	92.53
Bacteriophage Tuc2009	38.3	56	56	92.12
Bacteriophage Vf12	8.0	7	7	55.02
Bacteriophage Vf33	8.0	7	7	55.02
Bacteriophage VfO3K6	8.8	10	10	79.52
Bacteriophage VfO4K68	6.9	8	8	77.62
Bacteriophage VSKK	6.8	6	6	49.49
Bacteriophage VT2-Sa	60.9	83	86	88.79
Bacteriophage VWB	49.2	61	61	82.15
Bacteriophage WPhi	32.7	44	44	92.64
Bacteriophage X2	43.4	77	77	92.85
Bordetella phage BIP-1	42.6	48	48	95.65
Bordetella phage BMP-1	42.7	47	47	94.84
Bordetella phage BPP-1	42.5	49	49	96.37
Burkholderia cenocepacia phage	48.2	71	71	91.74

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
Bcep1				
Burkholderia cenocepacia phage BcepB1A	47.4	58	58	89.26
Burkholderia cenocepacia phage BcepMu	36.7	53	53	94.23
Burkholderia cepacia complex phage BcepC6B	42.4	46	46	92.67
Burkholderia cepacia phage Bcep22	63.9	78	79	91.44
Burkholderia cepacia phage Bcep43	48.0	65	66	92.81
Burkholderia cepacia phage Bcep781	48.2	66	67	92.83
Burkholderia cepacia phage BcepNazgul	58.1	75	75	90.26
Chlamydia phage 2	4.6	8	8	90.29
Chlamydia phage 3	4.6	8	8	89.09
Chlamydia pneumoniae phage CPAR39	4.5	7	7	87.51
Chlamydia psittaci bacteriophage chp1	4.9	12	12	68.59
Coliphage alpha3	6.1	10	10	83.77
Coliphage ID11	5.6	11	11	94.94
Coliphage phiK	6.1	10	10	83.84
Coliphage phiX174	5.4	11	11	95.08
Cyanophage P60	47.9	80	80	81.71
Cyanophage P-SSM2	252.4	329	330	94.73
Cyanophage P-SSM4	178.2	198	198	92.61
Cyanophage P-SSP7	45.0	53	53	88.76
Ehux virus	407.0	472	477	81.71
Enterobacteria phage 186	30.6	46	47	93.29
Enterobacteria phage epsilon15	39.7	49	49	92.39
Enterobacteria phage FI	4.3	4	4	95.42
Enterobacteria phage fr	3.6	4	4	90.80
Enterobacteria phage G4	5.6	11	11	94.94
Enterobacteria phage GA	3.5	3	3	91.32
Enterobacteria phage HK022	40.8	57	57	84.17
Enterobacteria phage HK97	39.7	61	61	88.50
Enterobacteria phage I2-2	6.7	9	9	66.73
Enterobacteria phage If1	8.5	10	10	80.18
Enterobacteria phage Ike	6.9	10	10	87.78
Enterobacteria phage KU1	3.5	4	4	91.80
Enterobacteria phage M13	6.4	10	10	83.28
Enterobacteria phage Mu	36.7	55	55	94.77

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
Enterobacteria phage N15	46.4	60	60	90.90
Enterobacteria phage P1	94.8	110	113	87.20
Enterobacteria phage P2	33.6	43	43	92.28
Enterobacteria phage P22	41.7	72	72	89.78
Enterobacteria phage PRD1	14.9	22	22	82.15
Enterobacteria phage RB43	180.5	292	293	93.80
Enterobacteria phage RB49	164.0	274	274	93.76
Enterobacteria phage RB69	167.6	273	275	93.83
Enterobacteria phage S13	5.4	12	12	96.47
Enterobacteria phage Sf6	39.0	66	68	90.30
Enterobacteria phage SP6	43.8	52	52	91.04
Enterobacteria phage T1	48.8	78	78	91.24
Enterobacteria phage T3	38.2	55	55	91.15
Enterobacteria phage T4	168.9	278	286	94.49
Enterobacteria phage T5	121.8	162	187	81.31
Enterobacteria phage T7	39.9	60	60	92.02
Enterobacterio phage MS2	3.6	4	4	90.95
Enterobacteriophage Qbeta	4.2	4	4	95.30
Guinea pig Chlamydia phage	4.5	9	9	85.03
Haemophilus phage HP1	32.4	42	43	92.33
Haemophilus phage HP2	31.5	37	38	90.88
Halorubrum phage HF2	77.7	114	120	88.36
Lactobacillus bacteriophage phi adh	43.8	63	63	90.68
Lactobacillus casei bacteriophage A2	43.4	61	62	88.69
Lactobacillus johnsonii prophage Lj928	38.4	50	51	90.46
Lactobacillus johnsonii prophage Lj965	40.2	46	50	89.58
Lactobacillus plantarum bacteriophage LP65	131.5	165	193	85.99
Lactobacillus plantarum bacteriophage phiJL-1	36.7	46	47	85.23
Lactococcus phage BK5-T	40.0	63	63	90.87
Lactococcus phage c2	22.2	39	39	92.60
Lactococcus phage P335	36.6	49	49	92.91
Lactococcus phage TP901-1	37.7	56	57	92.60
Lactococcus phage ul36	36.8	58	58	88.77
Listeria phage 2389	37.6	59	60	90.66
Listonella pelagia phage phiHSIC	38.0	47	47	87.65
Methanobacterium phage psiM2	26.1	32	32	87.51
Methanothermobacter wolfeii	28.8	35	35	88.65

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
prophage psiM100				
Mycobacteria phage D29	49.1	79	84	91.43
Mycobacteriophage PG1	69.0	100	100	94.25
Mycobacterium phage Barnyard	70.8	109	109	94.43
Mycobacterium phage Bxb1	50.6	86	86	90.41
Mycobacterium phage Bxz1	156.1	225	254	91.52
Mycobacterium phage Bxz2	50.9	86	89	91.20
Mycobacterium phage Che8	59.5	112	112	94.80
Mycobacterium phage Che9c	57.1	84	85	94.08
Mycobacterium phage Che9d	56.3	111	111	94.71
Mycobacterium phage Cjw1	75.9	141	143	93.03
Mycobacterium phage Corndog	69.8	122	122	93.88
Mycobacterium phage L5	52.3	85	88	88.13
Mycobacterium phage Omega	110.9	237	239	93.95
Mycobacterium phage Rosebush	67.5	90	90	95.26
Mycobacterium phage TM4	52.8	89	89	91.91
Mycoplasma arthritis bacteriophage MAV1	15.6	15	15	90.23
Propionibacterium phage phiB5	5.8	10	10	94.38
Pseudoalteromonas phage PM2	10.1	22	22	92.82
Pseudomonas aeruginosa bacteriophage PaP2	43.8	58	58	92.40
Pseudomonas aeruginosa phage F116	65.2	70	70	92.91
Pseudomonas aeruginosa phage PaP3	45.5	71	75	92.25
Pseudomonas bacteriophage phi-13	13.7	13	13	84.44
Pseudomonas phage D3	56.4	95	99	89.65
Pseudomonas phage gh-1	37.4	42	42	93.26
Pseudomonas phage Pf1	7.4	14	14	90.92
Pseudomonas phage Pf3	5.8	9	9	92.78
Pseudomonas phage phi-6	13.4	13	13	78.94
Pseudomonas phage phiKZ	280.3	306	313	87.27
Pseudomonas phage PP7	3.6	4	4	95.62
Ralstonia phage p12J	7.1	10	10	77.19
Roseophage SIO1	39.9	34	34	77.12
Salmonella typhimurium bacteriophage ES18	46.9	79	79	91.68
Salmonella typhimurium bacteriophage ST104	41.4	63	63	90.58
Salmonella typhimurium bacteriophage ST64T	40.7	65	66	90.77
Salmonella typhimurium phage	40.1	56	56	88.81

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
ST64B				
Shigella flexneri bacteriophage V	37.1	53	53	92.67
Sinorhizobium meliloti phage PBC5	57.4	83	83	89.76
Spiroplasma phage 1-C74	7.8	13	13	79.72
Spiroplasma phage 1-R8A2B	8.3	12	12	76.16
Spiroplasma phage 4	4.4	9	9	82.29
Staphylococcus aureus bacteriophage PVL	41.4	62	62	89.79
Staphylococcus aureus phage phi 11	43.6	53	53	79.90
Staphylococcus aureus phage phi 12	45.0	49	49	79.65
Staphylococcus aureus phage phi 13	42.7	49	49	78.07
Staphylococcus aureus phage phiP68	18.2	22	22	92.25
Staphylococcus aureus prophage phiPV83	45.6	65	65	87.22
Staphylococcus aureus temperate phage phiSLT	42.9	61	61	89.40
Staphylococcus phage 44AHJD	16.8	21	21	91.56
Staphylococcus phage K	127.4	115	119	78.68
Staphylococcus phage phiN315	44.1	65	65	87.93
Staphylococcus phage Twort	130.7	195	196	88.06
Streptococcus mitis phage SM1	34.7	56	56	91.43
Streptococcus phage C1	16.7	20	20	87.15
Streptococcus phage Cp-1	19.3	28	28	86.41
Streptococcus pneumoniae bacteriophage MM1	40.2	53	53	94.37
Streptococcus pyogenes phage 315.1	39.5	56	56	87.42
Streptococcus pyogenes phage 315.2	41.1	60	60	88.66
Streptococcus pyogenes phage 315.3	34.4	52	52	90.96
Streptococcus pyogenes phage 315.4	41.8	64	64	88.42
Streptococcus pyogenes phage 315.5	38.2	55	55	86.96
Streptococcus pyogenes phage 315.6	40.0	51	51	90.64
Streptococcus thermophilus bacteriophage 2972	34.7	44	44	93.35
Streptococcus thermophilus bacteriophage 7201	35.5	46	46	93.21
Streptococcus thermophilus bacteriophage DT1	34.8	45	45	89.18
Streptococcus thermophilus bacteriophage Sfi11	39.8	53	53	93.06
Streptococcus thermophilus bacteriophage Sfi19	37.4	45	45	88.93
Streptococcus thermophilus	40.7	50	50	89.36

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
bacteriophage Sfi21				
Streptococcus thermophilus temperate bacteriophage O1205	43.1	57	57	92.89
Stx1 converting bacteriophage	59.9	167	167	89.32
Stx2 converting bacteriophage I	61.8	166	169	89.44
Stx2 converting bacteriophage II	62.7	170	173	88.21
Temperate phage phiNIH1.1	41.8	55	55	82.01
Vibrio cholerae O139 fs1 phage	6.3	15	15	82.08
Vibrio cholerae phage KSF-1phi	7.1	12	12	73.11
Vibrio cholerae phage VGJphi	7.5	13	13	83.07
Vibrio harveyi bacteriophage VHML	43.2	57	57	83.29
Vibrio phage fs2	8.7	9	9	84.78
Vibrio phage VP5	39.8	48	48	93.08
Vibrio phage VSK	6.9	14	14	85.34
Vibriophage VP2	39.9	47	47	92.76
Vibriophage VP4	39.5	21	21	68.82
Vibriophage VpV262	46.0	67	68	91.79
Xanthomonas campestris pv. pelargonii phage Xp15	55.8	84	84	90.28
Xanthomonas oryzae bacteriophage Xp10	44.4	60	60	89.55
Xanthomonas phage Cf1c	7.3	9	9	80.32
Yersinia pestis phage phiA1122	37.6	50	50	91.84
Mitochondria				
<i>Acanthaster brevispinus</i> mitochondrion	16.3	13	37	
<i>Acanthaster planci</i> mitochondrion	16.2	13	37	
<i>Acrochordus granulatus</i> mitochondrion	17.6	13	37	64.00
<i>Alectura lathamii</i> mitochondrion	16.7	13	37	68.00
<i>Ambystoma andersoni</i> mitochondrion	16.4	13	37	69.00
<i>Ambystoma californiense</i> mitochondrion	16.4	13	37	69.00
<i>Ambystoma dumerilii</i> mitochondrion	16.4	13	37	69.00
<i>Andrias japonicus</i> mitochondrion	16.3	13	37	69.00
<i>Anisakis simplex</i> mitochondrion	13.9	12	36	
<i>Aplysia californica</i> mitochondrion	14.1	13	37	77.00
<i>Arabidopsis thaliana</i> mitochondrion	366.9	66	85	11.00
<i>Aspergillus niger</i> mitochondrion	31.1	16	43	
<i>Aspergillus tubingensis</i>	33.7	16	43	

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
mitochondrion				
<i>Asterias amurensis</i> mitochondrion	16.4	13	32	69.00
<i>Astropecten polyacanthus</i> mitochondrion	16.3	13	32	70.00
mitochondrion				
<i>Bactrocera oleae</i> mitochondrion	15.8	13	37	70.00
<i>Balaenoptera edeni</i> mitochondrion	16.4	13	37	
<i>Balaenoptera omurai</i> mitochondrion	16.4	13	37	
<i>Balaenoptera physalus</i> mitochondrion	16.4	13	37	69.00
<i>Bipes canaliculatus</i> mitochondrion	16.2	13	37	69.00
<i>Boa constrictor</i> mitochondrion	18.9	13	37	59.00
<i>Bombina bombina</i>	17.6	13	37	64.00
<i>Bombyx mandarina</i> mitochondrion	15.9	13	37	70.00
<i>Bombyx mori</i> mitochondrion	15.6	13	13	71.00
<i>Bos grunniens</i> mitochondrion	16.3	12	36	58.00
<i>Bos indicus</i> mitochondrion	16.3	13	36	69.00
<i>Bos taurus</i> mitochondrion	16.3	13	58	69.00
<i>Branta canadensis</i> mitochondrion	16.8	13	37	67.00
<i>Buergeria buergeri</i> mitochondrion	20.0	13	37	56.00
<i>Cafeteria roenbergensis</i> mitochondrion	43.1	34	58	82.00
mitochondrion				
<i>Callinectes sapidus</i> mitochondrion	16.3	13	37	68.00
<i>Campanulotes bidentatus compar</i> mitochondrion	14.8	13	37	
mitochondrion				
<i>Candida albicans SC5314</i> mitochondrion	40.4	13	45	24.00
mitochondrion				
<i>Candida glabrata</i> mitochondrion	20.0	11	37	47.00
<i>Candida metapsilosis</i> mitochondrion	24.2	15	41	54.00
<i>Candida orthopsilosis</i> mitochondrion	22.5	15	41	56.00
<i>Candida parapsilosis</i> mitochondrion	32.7	20	46	
<i>Carassius auratus x Cyprinus carpio</i> mitochondrion	16.6	13	37	
mitochondrion				
<i>Carassius auratus x Cyprinus carpio x Carassius cuvieri</i> mitochondrion	16.6	13	37	
<i>Cathartes aura</i> mitochondrion	16.8	13	37	67.00
<i>Centruroides limpidus</i> mitochondrion	14.5	13	36	74.00
<i>Cercopithecus aethiops</i> mitochondrion	16.4	13	37	69.00
mitochondrion				
<i>Cervus elaphus</i> mitochondrion	16.4	13	37	65.00
<i>Cervus nippon yakushimae</i> mitochondrion	16.4	13	37	68.00
mitochondrion				
<i>Chauliodus sloani</i> mitochondrion	17.8	13	37	64.00

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Chirocentrus dorab</i> mitochondrion	16.0	13	37	71.00
<i>Chlorurus sordidus</i> mitochondrion	16.7	13	37	68.00
<i>Choloepus didactylus</i> mitochondrion	16.5	13	37	68.00
<i>Chondrus crispus</i> mitochondrion	25.8	25	51	71.00
<i>Chrysodidymus synuroideus</i> mitochondrion	34.1	37	62	
<i>Ciona savignyi</i> mitochondrion	14.7	12	38	74.00
<i>Cobitis sinensis</i> mitochondrion	16.6	13	37	68.00
<i>Cordylus warreni</i> mitochondrion	17.2	13	39	66.00
<i>Coreoleuciscus splendidus</i> mitochondrion	16.6	13	37	68.00
<i>Crassostrea virginica</i> mitochondrion	17.2	12	37	
<i>Cricetulus griseus</i> mitochondrion	16.3	13	15	
<i>Crinipellis perniciosus</i> mitochondrion	109.1	89	117	50.00
<i>Crocidura russula</i> mitochondrion	17.2	13	26	65.00
<i>Cromeria nilotica</i> mitochondrion	16.0	13	35	
<i>Cyanidioschyzon merolae</i> mitochondrion	32.2	34	62	76.00
<i>Cyanidioschyzon merolae</i> mitochondrion	32.2	31	62	76.00
<i>Cygnus columbianus</i> mitochondrion	16.7	13	37	68.00
<i>Cylindrophis ruffus</i> mitochondrion	17.5	13	37	64.00
<i>Cypselurus hiraii</i> mitochondrion	16.5	13	37	69.00
<i>Danio rerio</i> mitochondrion	16.6	13	37	68.00
<i>Dasyurus hallucatus</i> mitochondrion	17.2	13	37	68.00
<i>Denticeps clupeoides</i> mitochondrion	17.2	13	37	65.00
<i>Desmarestia viridis</i> mitochondrion	39.0	39	68	69.00
<i>Dictyostelium citrinum</i> mitochondrion	58.6	33	46	68.00
<i>Dictyota dichotoma</i> mitochondrion	31.6	38	66	77.00
<i>Diomedea melanophris</i> mitochondrion	17.0	13	37	69.00
<i>Doliolum nationalis</i> mitochondrion	16.4	13	40	66.00
<i>Drosophila melanogaster</i> mitochondrion	19.5	13	37	57.00
<i>Echinococcus multilocularis</i> mitochondrion	13.7	12	36	73.00
<i>Echymipera rufescens australis</i> mitochondrion	16.6	13	37	68.00
<i>Emmelichthys struhsakeri</i> mitochondrion	16.5	13	37	68.00
<i>Epidermophyton floccosum</i> mitochondrion	30.9	64	51	67.00

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Epigonichthys lucayanus</i> mitochondrion	15.1	13	37	74.00
<i>Epigonichthys maldivensis</i> mitochondrion	15.0	13	37	75.00
<i>Fejervarya limnocharis</i> mitochondrion	17.7	13	38	63.00
<i>Fucus vesiculosus</i> mitochondrion	36.4	38	67	77.00
<i>Gallus gallus gallus</i> mitochondrion	16.8	13	37	67.00
<i>Gallus lafayettei</i> mitochondrion	16.8	13	37	67.00
<i>Gallus sonneratii</i> mitochondrion	16.8	13	37	67.00
<i>Gavia stellata</i> mitochondrion	17.6	13	37	67.00
<i>Gegeneophis ramaswamii</i> mitochondrion	15.9	13	36	70.00
<i>Gekko gecko</i> mitochondrion	16.4	13	37	68.00
<i>Geochelone pardalis</i> mitochondrion	19.4	13	37	54.00
<i>Geothelphusa dehaani</i> mitochondrion	18.2	13	38	61.00
<i>Gonodactylus chiragra</i> mitochondrion	16.3	13	37	68.00
<i>Grasseichthys gabonensis</i> mitochondrion	16.9	13	37	11.00
<i>Gryllotalpa orientalis</i> mitochondrion	15.5	13	37	71.00
<i>Gymnocrinus richeri</i> mitochondrion	16.0	13	37	71.00
<i>Haematobia irritans irritans</i> mitochondrion	16.1	13	37	69.00
<i>Haliotis rubra</i> mitochondrion	16.9	13	37	69.00
<i>Halocynthia roretzi</i> mitochondrion	14.8	12	37	72.00
<i>Hanseniopsis uvarum</i> mitochondrion	18.8	8	33	32.00
<i>Harpisquilla harpax</i> mitochondrion	15.7	13	37	72.00
<i>Hemibarbus labeo</i> mitochondrion	16.6	13	37	68.00
<i>Hemibarbus longirostris</i> mitochondrion	16.6	13	37	68.00
<i>Hemibarbus mylodon</i> mitochondrion	16.6	13	37	68.00
<i>Hexatrygon bickelli</i> mitochondrion	17.5	13	37	65.00
<i>Hyla chinensis</i> mitochondrion	18.2	13	37	62.00
<i>Ichthyophis bannanicus</i> mitochondrion	16.0	13	37	70.00
<i>Ilyanassa obsoleta</i> mitochondrion	15.3	13	37	73.00
<i>Indotestudo elongata</i> mitochondrion	16.8	13	37	67.00
<i>Japyx solifugus</i> mitochondrion	15.8	13	37	70.00
<i>Jenkinsia lamprotaenia</i> mitochondrion	16.8	13	37	67.00

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Kaloula pulchra</i> mitochondrion	16.8	13	37	67.00
<i>Kluyveromyces thermotolerans</i> mitochondrion	23.6	11	38	41.00
<i>Kneria</i> sp. <i>SL-2004</i> mitochondrion	16.4	13	37	43.00
<i>Larus dominicanus</i> mitochondrion	16.7	13	37	68.00
<i>Lates calcarifer</i> mitochondrion	16.5	13	37	70.00
<i>Latimeria menadoensis</i> mitochondrion	16.5	13	37	65.00
<i>Lecanicillium muscarium</i> mitochondrion	24.5	15	42	58.00
<i>Lepeophtheirus salmonis</i> mitochondrion	15.4	12	36	70.00
<i>Leptorhynchoides thecatus</i> mitochondrion	13.9	12	38	74.00
<i>Leptotrombidium akamushi</i> mitochondrion	13.7	13	37	75.00
<i>Leptotrombidium deliense</i> mitochondrion	13.7	13	37	74.00
<i>Leptotrombidium pallidum</i> mitochondrion	16.8	13	38	61.00
<i>Limnonectes fujianensis</i> mitochondrion	17.7	13	37	42.00
<i>Lipotes vexillifer</i> mitochondrion	16.4	13	37	69.00
<i>Lottia digitalis</i> mitochondrion	26.8	13	35	43.00
<i>Loxodonta africana</i> mitochondrion	16.9	13	37	67.00
<i>Luidia quinalia</i> mitochondrion	16.5	13	32	69.00
<i>Lyciasalamandra atifi</i> mitochondrion	16.7	13	37	68.00
<i>Lysiosquillina maculata</i> mitochondrion	16.3	13	37	68.00
<i>Macrobrachium rosenbergii</i> mitochondrion	15.8	13	37	68.00
<i>Malacochersus tornieri</i> mitochondrion	19.2	13	38	59.00
<i>Mammuthus primigenius</i> mitochondrion	16.8	13	37	67.00
<i>Manouria emys</i> mitochondrion	16.5	13	37	74.00
<i>Mantella madagascariensis</i> mitochondrion	22.9	13	38	74.00
<i>Marsupenaeus japonicus</i> mitochondrion	16.0	13	37	69.00
<i>Masturus lanceolatus</i> mitochondrion	16.5	13	37	70.00
<i>Melanogrammus aeglefinus</i>	16.6	13	37	68.00

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
mitochondrion				
<i>Menura novaehollandiae</i>	17.8	13	37	69.00
mitochondrion				
<i>Merlangius merlangus</i> mitochondrion	16.6	13	37	69.00
<i>Mesobuthus gibbosus</i> mitochondrion	16.0	13	37	67.00
<i>Microhyla heymonsi</i> mitochondrion	16.7	13	37	67.00
<i>Mola mola</i> mitochondrion	16.5	13	37	67.00
<i>Montastraea annularis</i> mitochondrion	16.1	13	17	72.00
<i>Montastraea faveolata</i> mitochondrion	16.1	13	17	72.00
<i>Mortierella verticillata</i>	58.7	25	54	36.00
mitochondrion				
<i>Mus musculus</i> mitochondrion	16.3	13	57	69.00
<i>Mytilus trossulus</i> mitochondrion	18.7	12	38	73.00
<i>Myxocyprinus asiaticus</i>	16.6	13	37	68.00
mitochondrion				
<i>Octopus ocellatus</i> mitochondrion	16.0	13	37	70.00
<i>Orbinia latreillii</i> mitochondrion	15.6	13	37	68.00
<i>Oreochromis mossambicus</i>	16.6	13	37	69.00
mitochondrion				
<i>Ornithodoros porcinus</i> mitochondrion	14.4	13	37	75.00
<i>Oryza sativa (indica cultivar-group)</i>	491.5	54	93	68.00
mitochondrion				
<i>Ovophis okinavensis</i> mitochondrion	17.4	13	37	64.00
<i>Pantholops hodgsonii</i> mitochondrion	16.5	13	37	68.00
<i>Paracoccidioides brasiliensis</i>	71.3	17	44	68.00
mitochondrion				
<i>Parakneria cameronensis</i>	16.6	13	37	66.00
mitochondrion				
<i>Paramecium aurelia</i> mitochondrion	40.5	46	54	64.00
<i>Paramesotriton hongkongensis</i>	16.3	13	37	69.00
mitochondrion				
<i>Pedinomonas minor</i> mitochondrion	25.1	11	23	69.00
<i>Periplaneta fuliginosa</i> mitochondrion	15.0	13	37	74.00
<i>Petrobius brevistylis</i> mitochondrion	15.7	13	37	71.00
<i>Phanogenia gracilis</i> mitochondrion	15.9	13	37	72.00
<i>Philaenus spumarius</i> mitochondrion	16.3	13	37	67.00
<i>Phocoena phocoena</i> mitochondrion	16.4	13	26	59.00
<i>Phractolaemus ansorgii</i>	16.5	13	37	64.00
mitochondrion				
<i>Physcomitrella patens</i> mitochondrion	105.3		69	68.00
<i>Phytophthora infestans</i>	38.0	40	67	69.00
mitochondrion				

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Placopecten magellanicus</i> mitochondrion	32.1	12	48	35.00
<i>Plasmodium simium</i> mitochondrion	6.0	3	3	55.00
<i>Plasmodium vivax</i> mitochondrion	6.0	3	3	56.00
<i>Podospora anserina</i> mitochondrion	94.2	53	82	58.00
<i>Polypedates megacephalus</i> mitochondrion	16.5	11	35	56.00
<i>Porichthys myriaster</i> mitochondrion	18.9	13	37	60.00
<i>Pseudocarcinus gigas</i> mitochondrion	15.5	13	37	67.00
<i>Pterodroma brevirostris</i> mitochondrion	16.4	13	37	67.00
<i>Pteronarcys princeps</i> mitochondrion	16.0	13	37	70.00
<i>Pyrocoelia rufa</i> mitochondrion	17.7	13	37	62.00
<i>Python regius</i> mitochondrion	17.2	13	37	65.00
<i>Raja porosa</i> mitochondrion	17.0	13	37	67.00
<i>Rangifer tarandus</i> mitochondrion	16.4	13	37	69.00
<i>Ranzania laevis</i> mitochondrion	16.5	13	37	69.00
<i>Rattus norvegicus</i> mitochondrion	16.3	13	37	69.00
<i>Rattus norvegicus strain BN/SsNHsdMCW</i> mitochondrion	16.3	13	37	69.00
<i>Reclinomonas americana</i> mitochondrion	69.0	67	97	80.00
<i>Rhacophorus schlegelii</i> mitochondrion	21.4	13	37	52.00
<i>Rhizopus oryzae</i> mitochondrion	54.2	24	51	38.00
<i>Rhodeus uyekii</i> mitochondrion	16.8	13	37	69.00
<i>Roboastra europaea</i> mitochondrion	14.5	13	37	75.00
<i>Rousettus aegyptiacus</i> mitochondrion	16.7	13	37	67.00
<i>Saccharomyces cerevisiae</i> mitochondrion	85.8	19	46	23.00
<i>Saccoglossus kowalevskii</i> mitochondrion	17.0	13	37	66.00
<i>Salanx ariakensis</i> mitochondrion	16.6	13	37	68.00
<i>Scenedesmus obliquus</i> mitochondrion	42.8	20	53	38.00
<i>Schizosaccharomyces pombe</i> mitochondrion	19.4	10	37	50.00
<i>Scleropages formosus</i> mitochondrion	16.7	13	37	68.00
<i>Sclerophasma paresisensis</i> mitochondrion	15.5	13	37	71.00
<i>Scolecophorus vittatus</i> mitochondrion	16.0	13	37	70.00
<i>Scomber scombrus</i> mitochondrion	16.6	13	37	68.00

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Sebastes schlegeli</i> mitochondrion	16.5	13	37	69.00
<i>Sepia officinalis</i> mitochondrion	16.2	13	37	65.00
<i>Sepioteuthis lessoniana</i> mitochondrion	16.6	13	37	69.00
<i>Siphonops annulatus</i> mitochondrion	16.2	13	37	69.00
<i>Sminthopsis crassicaudata</i> mitochondrion	17.4	13	37	69.00
<i>Smittium culisetae</i> mitochondrion	58.7	31	61	43.00
<i>Sorex unguiculatus</i> mitochondrion	17.1	13	37	66.00
<i>Spadella cephaloptera</i> mitochondrion	11.9	11	13	81.00
<i>Spizaetus alboniger</i> mitochondrion	18.0	13	37	63.00
<i>Spizaetus nipalensis</i> mitochondrion	17.7	13	37	64.00
<i>Squilla empusa</i> mitochondrion	15.8	13	37	70.00
<i>Steinernema carpocapsae</i> mitochondrion	13.9	12	36	73.00
<i>Sundasalanx mekongensis</i> mitochondrion	16.7	13	37	68.00
<i>Sus scrofa</i> mitochondrion	16.6	13	37	68.00
<i>Synodus variegatus</i> mitochondrion	16.5	13	37	69.00
<i>Taenia asiatica</i> mitochondrion	13.7	12	36	73.00
<i>Taeniopygia guttata</i> mitochondrion	16.9	13	37	81.00
<i>Tamolanica tamolana</i> mitochondrion	16.1	13	37	69.00
<i>Teratoscincus keyserlingii</i> mitochondrion	17.2	13	37	64.00
<i>Testudo graeca</i> mitochondrion	19.3	13	37	74.00
<i>Testudo horsfieldii</i> mitochondrion	17.1	13	37	66.00
<i>Testudo kleinmanni</i> mitochondrion	17.7	13	37	64.00
<i>Testudo marginata</i> mitochondrion	19.5	13	37	58.00
<i>Tetrahymena pyriformis</i> mitochondrion	47.3	44	58	62.00
<i>Tetraodon nigroviridis</i> mitochondrion	16.5	13	37	69.00
<i>Thalassiosira pseudonana</i> mitochondrion	43.8	35	62	64.00
<i>Thrips imaginis</i> mitochondrion	15.4	13	38	70.00
<i>Triticum aestivum</i> mitochondrion	452.5	39	73	7.00
<i>Verasper variegatus</i> mitochondrion	17.3	13	37	24.00
<i>Verasper variegatus</i> mitochondrion	17.0	13	37	67.00
<i>Verasper variegatus</i> mitochondrion	14.8	13	39	74.00
<i>Volemys kikuchii</i> mitochondrion	16.3	13	37	69.00
<i>Watasenia scintillans</i> mitochondrion	20.1	18	43	
<i>Xenopeltis unicolor</i> mitochondrion	18.9	13	37	59.00

Species or strain	Genome size (kbp)	Protein-coding gene number	Total gene number*	Gene-coding percentage**
<i>Yarrowia lipolytica</i> mitochondrion	48.0	24	53	45.00
Chloroplasts				
<i>Acorus calamus</i> plastid	153.8	84	130	49.00
<i>Adiantum capillus-veneris</i> chloroplast	151.0	87	130	
<i>Anthoceros formosae</i> chloroplast	161.0	90	135	
<i>Arabidopsis thaliana</i> chloroplast	154.0	87	132	
<i>Atropa belladonna</i> chloroplast	156.7	87	132	51.00
<i>Calycanthus floridus</i> var. <i>glaucus</i> chloroplast	153.0	88	134	
<i>Chaetosphaeridium globosum</i> chloroplast	131.0	98	141	
<i>Cucumis sativus</i> chloroplast	155.3	85	131	49.00
<i>Cyanophora paradoxa</i> chloroplast	135.6	149	192	
<i>Emiliana huxleyi</i> chloroplast	105.3	119	155	76.00
<i>Euglena gracilis</i> chloroplast	143.0	66	114	66.00
<i>Glycine max</i> chloroplast	152.0		128	48.00
<i>Gossypium hirsutum</i> chloroplast	160.0		131	44.00
<i>Gracilaria tenuistipitata</i> var. <i>liui</i> chloroplast	183.9	203	238	80.00
<i>Lactuca sativa</i> chloroplast	152.7	84	128	47.00
<i>Lycopersicon esculentum</i> chloroplast	155.0	86	132	54.00
<i>Nicotiana tabacum</i> plastid	155.9	101	146	49.00
<i>Nicotiana tomentosiformis</i> chloroplast	155.7	109	155	54.00
<i>Phalaenopsis aphrodite</i> subsp. <i>formosana</i> chloroplast	149.0	97	141	44.00
<i>Porphyra yezoensis</i> chloroplast	192.0	209	264	51.00
<i>Psilotum nudum</i> chloroplast	138.8	101	150	55.00
<i>Solanum bulbocastanum</i> chloroplast	155.0		142	80.00
<i>Spinacia oleracea</i> plastid	151.0	98	143	48.00
<i>Theileria parva</i> strain Muguga apicoplast	39.6	44	70	66.00
<i>Toxoplasma gondii</i> apicoplast chloroplast	35.0	26	63	52.00

*Total gene number equal to protein-coding gene number plus non-protein coding RNA gene number. Sources of data are as shown in the reference list below.

**Most gene-coding percentage data were retrieved from the percent of DNA coding number of bases over total number of bases from the Integrated Microbial Genomes (IMG) database in the Department of Energy Joint Genome Institute (JGI) website (<http://img.jgi.doe.gov/cgi-bin/pub/main.cgi>). Note that JGI's estimated DNA coding number of bases includes bases from introns and other untranslated regions in addition to coding regions, thus the gene-coding percentages were overestimated for most eukaryotes. Therefore, for eukaryotic gene-coding

percentages only data retrieved from peer-reviewed literature (in parentheses) were used in our analyses.

References:

1. Markowitz VM, Szeto E, Palaniappan K, Grechkin Y, Chu K, et al. (2007) The integrated microbial genomes (IMG) system in 2007: data content and analysis tool extensions. *Nucl Acids Res*: gkm846.
2. Reference Sequence (RefSeq) collection in National Center for Biotechnology Information (NCBI, <http://www.ncbi.nlm.nih.gov>).
3. Xu L, Chen H, Hu X, Zhang R, Zhang Z, et al. (2006) Average gene length is highly conserved in prokaryotes and eukaryotes and diverges only between the two kingdoms. *Mol Biol Evol* 23: 1107-1108.
4. Nierman WC, Pain A, Anderson MJ, Wortman JR, Kim HS, et al. (2005) Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. 438: 1151-1156.
5. Archibald, JM (2007) Nucleomorph genomes: structure, function, origin and evolution. *BioEssays* 29: 392-402.
6. Biology analysis group, Xia Q, Zhou Z, Lu C, Cheng D, et al. (2004) A draft sequence for the genome of the domesticated silkworm (*Bombyx mori*). *Science* 306: 1937-1940.
7. Merchant SS, Prochnik SE, Vallon O, Harris EH, Karpowicz SJ, et al. (2007) The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* 318: 245-250.
8. Lane CE, van den Heuvel K, Kozera C, Curtis BA, Parsons BJ, et al. (2007) Nucleomorph genome of *Hemiselmis andersenii* reveals complete intron loss and compaction as a driver of protein structure and function. *Proc. Natl. Acad. Sci. USA* 104: 19908-19913.
9. Matsuzaki M, Misumi O, Shin-i T, Maruyama S, Takahara M, et al. (2004) Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. *Nature* 428: 653-657.
10. International Human Genome Sequencing Consortium. (2004) Finishing the euchromatic sequence of the human genome. *Nature* 431: 931-945.
11. Derelle E, Ferraz C, Rombauts S, Rouzé P, Worden AZ, et al. (2006) Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proc. Natl. Acad. Sci. USA* 103: 11647-11652.
12. Aury JM, Jaillon O, Duret L, Noel B, Jubin C, et al. (2006) Global trends of whole-genome duplications revealed by the ciliate *Paramecium tetraurelia*. *Nature* 444: 171-178.
13. Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, et al. (2008) The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature* 456: 239-244.
14. Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, et al. (2004) The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science* 306: 79-86.
15. Eisen JA, Coyne RS, Wu M, Wu D, Thiagarajan M, et al. (2006) Macronuclear genome sequence of the ciliate *Tetrahymena thermophila*, a model eukaryote. *PLoS Biol* 4: e286.