

Table S3. Distribution of CDSs in the functional categories of the COG database presented as percentages to total CDS numbers.*

Category		Archaea**	Bacteria**	<i>Pth</i>	<i>Chy</i>	<i>Moth</i>	<i>Tte</i>	<i>Swol</i>	<i>Cth</i>	<i>Cac</i>	<i>Lla</i>	<i>Lpl</i>	<i>Efa</i>	<i>Bsu</i>	<i>Wsu</i>	<i>Sac</i>	<i>Sfu</i>	<i>Dde</i>	<i>Gsu</i>	<i>Zmo</i>	<i>Pae</i>	<i>Tde</i>	<i>Pac</i>	<i>Bfr</i>	<i>Tma</i>	<i>Msth</i>	<i>Mtht</i>	<i>Mhu</i>	<i>Mma</i>	<i>Mac</i>
J	Translation	4.98	3.66	4.83	4.03	4.19	4.36	5.94	5.25	3.47	4.64	3.70	3.51	3.00	4.20	3.70	4.09	3.06	3.54	5.13	2.60	4.07	4.65	2.79	5.21	8.71	5.86	4.76	6.64	2.91
A	RNA processing and modification	0.03	0.01	0.00	0.00	0.00	0.00	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.02	0.00	0.04	0.00	0.00	0.06	0.05	0.03	0.05	0.02
K	Transcription	3.50	4.95	4.18	3.89	5.64	5.24	5.51	5.68	6.40	5.40	8.14	6.27	6.34	3.58	2.89	3.83	3.52	3.60	3.48	6.74	3.36	5.45	3.52	3.89	4.04	3.21	4.14	3.59	3.16
L	Replication, recombination and repair	4.71	4.43	5.82	3.96	4.71	5.53	7.80	8.59	3.37	6.72	4.22	4.53	3.03	3.49	4.52	3.49	3.35	3.52	3.66	2.15	3.25	4.17	3.19	4.38	4.56	3.72	5.50	3.00	5.75
B	Chromatin structure and dynamics	0.10	0.00	0.10	0.00	0.00	0.04	0.12	0.03	0.02	0.00	0.00	0.00	0.00	0.00	0.03	0.10	0.02	0.03	0.00	0.00	0.00	0.00	0.05	0.40	0.20	0.25	0.22	0.04	
D	Cell cycle control, mitosis and meiosis	0.68	0.69	1.06	1.17	1.52	1.30	1.63	1.36	0.87	0.60	0.61	0.60	0.71	0.88	0.70	0.85	0.70	0.85	0.87	0.48	0.68	0.52	0.67	1.17	0.63	0.56	0.74	0.87	0.61
V	Defense mechanisms	1.02	1.15	0.89	0.88	0.93	1.55	2.10	0.96	2.38	1.56	1.59	1.47	1.16	1.19	1.22	1.44	1.14	0.98	1.05	1.08	3.32	1.60	2.32	1.51	0.87	0.71	1.30	0.71	1.52
T	Signal transduction mechanisms	1.92	3.45	2.88	3.57	3.60	4.26	5.04	5.50	4.51	1.88	2.66	2.70	3.47	6.94	3.85	6.70	6.13	7.54	2.47	4.80	3.05	2.28	3.68	2.87	1.85	1.94	8.06	1.58	2.19
M	Cell wall/membrane biogenesis	2.34	4.11	4.52	3.11	3.67	3.94	4.46	5.56	4.81	3.96	4.44	3.63	4.23	4.77	5.48	5.53	4.69	5.00	5.08	3.98	3.83	2.85	6.82	3.41	3.40	3.01	3.00	1.80	2.13
N	Cell motility	0.83	1.38	2.53	2.05	2.12	2.15	2.99	3.00	2.60	0.28	0.15	0.36	1.22	3.85	2.24	0.95	2.38	3.39	1.97	2.39	2.27	0.00	0.02	2.68	0.35	0.10	2.81	1.14	0.79
Z	Cytoskeleton	0.00	0.01	0.00	0.00	0.00	0.00	0.04	0.03	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W	Extracellular structures	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02
U	Intracellular trafficking and secretion	0.85	1.72	0.79	1.45	1.71	1.41	1.71	1.88	1.02	0.92	0.64	0.93	0.93	2.48	2.24	2.32	1.29	2.22	2.01	2.55	1.05	0.76	1.49	1.75	1.15	0.76	1.64	0.93	0.97
O	Posttranslational modification, protein turnover, chaperones	2.62	2.62	2.29	2.69	2.86	2.43	3.34	2.97	1.83	1.64	1.56	1.41	2.16	3.54	2.74	3.58	2.72	3.12	3.30	2.86	2.17	2.44	1.72	2.48	3.69	2.80	3.37	2.51	2.25
C	Energy production and conversion	5.68	4.37	6.13	6.89	7.20	4.26	5.24	3.37	2.90	2.80	3.33	2.91	3.98	7.03	5.57	9.16	5.25	6.27	3.48	4.99	2.31	4.85	3.23	5.21	7.21	7.69	6.52	7.95	5.33
G	Carbohydrate transport and metabolism	2.25	3.49	2.12	1.45	3.30	4.58	2.60	4.42	4.41	4.60	6.06	5.37	4.58	1.55	2.42	3.49	2.53	1.93	2.43	2.36	2.65	7.29	4.30	7.06	3.23	2.14	2.29	2.34	1.56
E	Amino acid transport and metabolism	5.57	5.91	5.79	6.43	6.75	6.97	5.70	5.25	5.40	6.49	6.12	5.16	6.52	6.37	4.31	6.63	6.03	4.81	6.32	7.22	4.27	6.41	3.97	8.47	6.52	5.04	5.56	6.21	5.08
F	Nucleotide transport and metabolism	1.89	1.63	2.09	1.84	1.89	1.87	2.10	1.98	1.61	2.72	2.39	2.55	1.67	2.08	1.60	1.63	1.19	1.35	2.15	1.57	1.53	2.65	1.36	2.39	2.89	2.50	1.85	2.50	1.24
H	Coenzyme transport and metabolism	3.40	2.65	3.84	3.85	4.34	2.11	4.23	3.15	2.78	2.56	2.45	1.92	2.54	3.63	2.77	4.22	2.77	2.70	3.80	2.58	1.87	4.13	2.65	2.53	6.75	4.38	5.10	5.83	2.89
I	Lipid transport and metabolism	1.67	2.66	2.47	2.61	1.34	1.80	2.60	1.14	1.78	2.32	2.26	1.80	2.54	1.50	1.95	2.17	1.09	1.64	2.34	3.79	1.39	2.00	1.40	1.31	1.27	1.27	0.93	0.87	0.77
P	Inorganic ion transport and metabolism	3.56	3.63	2.84	2.65	3.45	3.24	3.22	2.87	2.80	3.56	3.46	3.45	3.80	5.31	2.89	3.73	2.99	3.17	3.53	4.57	3.66	4.05	4.61	5.06	5.65	3.41	4.63	4.19	5.33
Q	Secondary metabolites synthesis, transport and catabolism	1.10	1.79	0.79	1.41	0.96	1.09	0.93	0.71	1.04	1.12	1.10	0.96	2.22	0.97	1.11	1.17	0.88	1.06	1.37	2.74	0.37	0.96	0.65	0.83	0.63	0.25	0.22	0.44	1.92
R	General function prediction only	11.30	8.19	7.81	7.81	7.98	8.48	8.65	9.02	8.80	9.13	9.51	8.07	8.65	7.65	7.87	10.33	6.98	7.75	8.47	9.27	7.22	7.69	6.91	10.03	11.19	12.33	11.80	11.06	11.68
S	Function unknown	6.43	4.81	5.89	5.16	5.86	4.54	9.00	6.70	5.88	5.52	4.83	5.13	5.54	4.02	4.34	5.48	3.82	4.52	3.57	5.87	5.12	3.77	3.23	4.72	10.33	9.37	9.54	9.69	6.94
-	Not categorized	33.56	32.69	30.34	33.11	25.97	28.86	23.90	20.57	31.28	31.55	30.77	37.28	31.70	24.98	35.56	24.49	37.46	30.97	33.52	25.38	42.56	31.42	41.49	22.98	24.93	28.68	16.00	25.87	34.89

*Refer to Table 1 for organism names.

**Average values for 34 bacteria and 12 archaea cited from reference 10 in the main text.