Supporting Information

Dutton *et al.* 10.1073/pnas.0804621105

SI Methods

We analyzed the protein sequences of these genomes with Phobius, http://phobius.sbc.su.se (1) and a Prosite profile for lipoproteins, release 20.0, www.expasy.ch/prosite (2). Phobius is a subcellular localization prediction program based on SignalP 3.0 and TMHMM 2.0. Thus, proteins exported by the general secretion machinery, SecYEG, should be detected, as well as many proteins that are exported by the major alternative pathway for export in many bacteria, the TAT pathway (3), which utilizes signal sequences that are very similar to Sec pathway signal sequences. The Sec system is universally conserved and signals in secreted and transmembrane proteins that determine secretion and topology are similar across all bacteria. Such signals, as identified by the methods we used, are variable but always a significant fraction of all of the proteins in each bacterial genome. Thus we believe our approach is adequate for estimation of gross statistical features of the distribution of cysteine residues in exported proteins of most, if not all organisms.

For each protein, each amino acid residue was assigned to one of six classes, based on its predicted subcellular localization. Thus, each amino acid within a protein was assigned to one of the following classes: cytoplasmic (class 1); transmembrane protein-cytoplasmic domains (class 2); transmembrane protein-inner membrane spanning helices (class 3); transmembrane protein-periplasmic domains (class 4); exported protein, directed by a signal sequence whether the final destination is the periplasm, the outer membrane or outside of the cell (class 5); and other, which includes residues predicted to be in cleavable signal sequences and the amino terminal cysteine residues of mature lipoproteins (class 6). Transmembrane proteins with predicted signal sequences were classified as transmembrane.

For each genome, we calculated two numbers for each of the twenty amino acids in each class. The first is the even fraction, the fraction of proteins with even numbers of that amino acid of that class, excluding proteins with none of that amino acid in that class. We term that number the even fraction, or Efrac. The second number, the AApref, is a measure of the preference for or bias against that amino acid in that class. This is calculated from the amino acid composition of the class and the amino acid composition of the whole proteome. It is the ratio of the frequency of the amino acid in the class to the frequency in the genome. This is the same as the ratio of the fraction of the amino acid that is in the class to the fraction of all amino acids that is in the class.

To assess the significance of the Efrac, we carried out a randomization procedure to obtain a mean value and standard deviation for the Efrac of each amino acid expected at random. We used two different randomization procedures for *E. coli*. In the first, we simply randomized the sequences of each class, keeping the overall amino acid composition of the class constant. By repeating this procedure 1,000 times and averaging the even fractions, we obtained mean and standard deviation values. Repetition of the entire process produced identical or nearly identical results. In the second method, we used random numbers generated according to the Poisson distribution to get counts for each protein. The Poisson parameter, lambda, was set to the number of amino acids of that protein in that class times the frequency of the amino acid in the class. Again repetition for 1000 times and averaging gave a mean and standard deviation which was use to calculate a z score, the number of standard deviations between the random mean and the observed value of the Efrac. We used a Perl interface to the C library, RANDLIB, obtained from Comprehensive Perl Archive Network, www.cpan.org. This method gave the same result as that described above for *E. coli* and was used to all other genomes since it is computationally faster.

DsbA homologs with a cytoplasmic localization, based on Phobius predictions, were excluded. Since the Pfam DsbB HMM model missed some known DsbB homologs found in the alpha-proteobacteria, we built an additional DsbB HMM model (based on alpha-proteobacterial DsbB sequences) to supplement the homology searches. We also used BLASTP (4) to identify additional DsbA homologs using the *Staphylococcus aureus* DsbA (gi11935158) as a query and collected hits below the *e* value of $<10-4$.

1. Kall L, Krogh A, Sonnhammer EL (2007) Advantages of combined transmembrane topology and signal peptide prediction—The Phobius web server. *Nucleic Acids Res* 35:W429–W432.
2. Hulo N, *et al.* (2006) The PROSITE database. *Nucleic Acids Res* 34:D227–D230.
3. Lee PA, Tullman-Ercek D, Georgiou G (2006) The bacterial twin-arginine translocation pathway. *Annu Rev Microbiol* 60:373–395.
4. Altschul SF, *et al.* (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410.

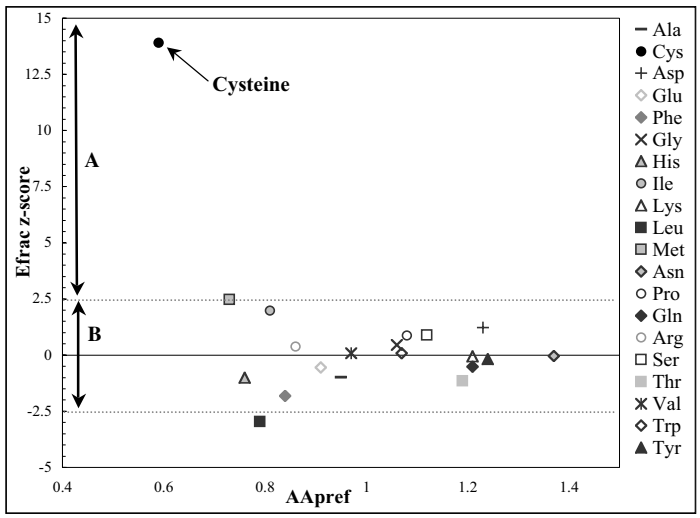

Fig. S1. Counting of all amino acids in *E. coli* K12 exported proteins. The z-score for the fraction of exported proteins with even numbers of an amino acid (Efrac), is plotted against the AApref for each amino acid (an AApref <1.0 indicates a bias against incorporation of the amino acid into exported proteins). Region A: There are significantly more even numbers of the amino acid in exported proteins than is predicted by the random model. Region B: Exported proteins do not have a significant bias for even numbers of these amino acids ($2.57 > z > -2.57$).

Phylum	Class	Organism Name	DsbA	DsbB	VKOR	Efrac z-score	Cpref	Efrac	Random Efrac mean	Oxy. Req	Habitat
Proteobacteria	Gamma	Acinetobacter_sp_ADP1	+	+	-	11.994	0.55	0.718	0.388 +/- 0.027	Aerobic	Multiple
		Alcanivorax_borkumensis_SK2	+	+	-	10.302	0.7	0.676	0.418 +/- 0.025	Aerobic	Aquatic
		Alkalilimnicola_ehrlichei_MLHE-1	+	+	-	10.322	0.6	0.692	0.39 +/- 0.029	Facultative	Aquatic
		Baumannia_cicadellinicola_Homalodisca_coagulata	+	-	-	-1.541	0.71	0.291	0.434 +/- 0.093		Host-associated
		Buchnera_aphidicola	-	-	-	1.166	1.21	0.625	0.478 +/- 0.126		Host-associated
		Buchnera_aphidicola_Cc_Cinara_cedri	-	-	-	0.805	0.92	0.625	0.488 +/- 0.17		Host-associated
		Buchnera_aphidicola_Sg	+	-	-	1.402	0.78	0.59	0.446 +/- 0.102		Host-associated
		Buchnera_sp	+	-	-	0.264	0.91	0.476	0.449 +/- 0.104		Host-associated
		Candidatus_Blochmannia_floridanus	-	+	-	0.273	1.02	0.5	0.466 +/- 0.124		Specialized
		Candidatus_Blochmannia_pennsylvanicus_BPEN	+	+	-	-1.507	0.87	0.32	0.461 +/- 0.094		
		Chromohalobacter_salexigens_DSM_3043	+	+	-	11.485	0.48	0.707	0.367 +/- 0.03	Aerobic	
		Colwellia_psychrerythraea_34H	+	+	-	17.286	0.5	0.73	0.392 +/- 0.02	Facultative	Specialized
		Coxiella_burnetii	+	+	-	2.095	0.86	0.505	0.423 +/- 0.039	Facultative	Multiple
		Erwinia_carotovora_atroseptica_SCRI1043	+	+	-	15.868	0.57	0.75	0.394 +/- 0.022	Facultative	Multiple
		Escherichia_coli_536	+	+	-	15.159	0.55	0.731	0.397 +/- 0.022	Facultative	Host-associated
		Escherichia_coli_CFT073	+	+	-	12.658	0.58	0.669	0.408 +/- 0.021	Facultative	Host-associated
		Escherichia_coli_K12	+	+	-	**14.041**	**0.59**	**0.714**	**0.402 +/- 0.022**	**Facultative**	**Host-associated**
		Escherichia_coli_O157H7	+	+	-	13.209	0.63	0.669	0.404 +/- 0.02		
		Escherichia_coli_O157H7_EDL933	+	+	-	13.769	0.62	0.7	0.403 +/- 0.022	Facultative	Host-associated
		Escherichia_coli_UTI89	+	+	-	14.233	0.6	0.702	0.409 +/- 0.021	Facultative	Host-associated
		Escherichia_coli_W3110	+	+	-	13.882	0.6	0.717	0.402 +/- 0.023	Facultative	Host-associated
		Francisella_tularensis_FSC_198	+	+	-	6.202	0.68	0.647	0.402 +/- 0.039		
		Francisella_tularensis_holarctica	+	+	-	5.402	0.67	0.621	0.403 +/- 0.04	Aerobic	Host-associated
		Francisella_tularensis_holarctica_OSU18	+	+	-	5.452	0.7	0.635	0.406 +/- 0.042	Aerobic	Multiple
		Francisella_tularensis_tularensis	+	+	-	6.14	0.68	0.653	0.404 +/- 0.041	Aerobic	Aquatic
		Haemophilus_ducreyi_35000HP	+	+	-	6.978	0.68	0.694	0.408 +/- 0.041	Anaerobic	Host-associated
		Haemophilus_influenzae	+	+	-	7.124	0.63	0.701	0.409 +/- 0.041	Facultative	Host-associated
		Haemophilus_influenzae_86_028NP	+	+	-	8.714	0.61	0.739	0.399 +/- 0.039	Facultative	Host-associated
		Haemophilus_somnus_129PT	+	+	-	7.57	0.71	0.714	0.414 +/- 0.04	Facultative	Host-associated
		Hahella_chejuensis_KCTC_2396	+	+	+	16.662	0.65	0.7	0.419 +/- 0.017	Facultative	Aquatic
		Idiomarina_loihiensis_L2TR	+	+	-	14.3	0.5	0.77	0.377 +/- 0.027	Aerobic	Specialized
		Legionella_pneumophila_Lens	+	+	-	8.05	0.84	0.645	0.439 +/- 0.026	Aerobic	Host-associated
		Legionella_pneumophila_Paris	+	+	-	8.684	0.84	0.648	0.438 +/- 0.024	Aerobic	Host-associated
		Legionella_pneumophila_Philadelphia_1	+	-	-	7.596	0.76	0.636	0.433 +/- 0.027	Aerobic	Host-associated
		Mannheimia_succiniciproducens_MBEL55E	+	+	-	7.626	0.59	0.662	0.4 +/- 0.034	Anaerobic	Host-associated
		Methylococcus_capsulatus_Bath	+	+	-	6.658	0.8	0.608	0.431 +/- 0.027	Aerobic	Multiple
		Nitrosococcus_oceani_ATCC_19707	+	+	-	10.356	0.65	0.683	0.402 +/- 0.027		Aquatic
		Pasteurella_multocida	+	+	-	9.755	0.52	0.729	0.399 +/- 0.034	Facultative	Host-associated
		Photobacterium_profundum_SS9	+	+	-	14.753	0.6	0.703	0.409 +/- 0.02	Facultative	Multiple
		Photorhabdus_luminescens	+	+	-	11.62	0.62	0.713	0.409 +/- 0.026	Facultative	Host-associated
		Pseudoalteromonas_atlantica_T6c	+	+	-	15.047	0.51	0.693	0.401 +/- 0.019	Aerobic	Aquatic
		Pseudoalteromonas_haloplanktis_TAC125	+	+	-	15.775	0.47	0.767	0.385 +/- 0.024	Aerobic	Aquatic
		Pseudomonas_aeruginosa	+	+	-	19.617	0.6	0.739	0.4 +/- 0.017	Aerobic	Multiple
		Pseudomonas_aeruginosa_UCBPP-PA14	+	+	-	19.087	0.61	0.735	0.403 +/- 0.017	Aerobic	Multiple
		Pseudomonas_entomophila_L48	+	+	-	15.835	0.59	0.697	0.398 +/- 0.019		Multiple
		Pseudomonas_fluorescens_Pf-5	+	+	-	15.843	0.63	0.697	0.411 +/- 0.018	Aerobic	Multiple
		Pseudomonas_fluorescens_PfO-1	+	+	-	17.318	0.61	0.706	0.392 +/- 0.018	Aerobic	Multiple
		Pseudomonas_putida_KT2440	+	+	-	13.904	0.57	0.676	0.399 +/- 0.02	Aerobic	Multiple
		Pseudomonas_syringae_phaseolicola_1448A	+	+	-	12.133	0.61	0.645	0.397 +/- 0.02	Aerobic	Multiple

Table S1. Complete dataset including homology data, cysteine counting data, and organismal information for all 375 genomes analyzed.

	Pseudomonas_syringae_pv_B728a	+	+	-	15.839	0.62	0.71	0.398 +/- 0.02	Aerobic	Multiple
	Pseudomonas_syringae_tomato_DC3000	+	+	-	14.016	0.61	0.674	0.396 +/- 0.02	Aerobic	Multiple
	Psychrobacter_arcticum_273-4	+	+	-	8.249	0.7	0.704	0.395 +/- 0.037		Specialized
	Psychrobacter_cryohalolentis_K5	+	+	-	10.224	0.66	0.742	0.39 +/- 0.034		Multiple
	Saccharophagus_degradans_2-40	+	+	+	15.819	0.65	0.718	0.43 +/- 0.018	Aerobic	Aquatic
	Salmonella_enterica_Choleraesuis	+	+	-	14.392	0.66	0.713	0.41 +/- 0.021		
	Salmonella_enterica_Paratypi_ATCC_9150	+	+	-	14.322	0.65	0.722	0.41 +/- 0.022		
	Salmonella_typhi	+	+	-	13.354	0.67	0.7	0.41 +/- 0.022	Facultative	Host-associated
	Salmonella_typhi_Ty2	+	+	-	13.849	0.67	0.701	0.409 +/- 0.021	Facultative	Host-associated
	Salmonella_typhimurium_LT2	+	+	-	14.502	0.66	0.73	0.415 +/- 0.022	Facultative	Host-associated
	Shewanella_denitrificans_OS217	+	+	-	15.605	0.6	0.744	0.406 +/- 0.022	Facultative	Aquatic
	Shewanella_frigidimarina_NCIMB_400	+	+	-	12.77	0.67	0.687	0.421 +/- 0.021	Facultative	Multiple
	Shewanella_MR-4	+	+	-	15.405	0.68	0.736	0.423 +/- 0.02	Facultative	Multiple
	Shewanella_MR-7	+	+	-	15.474	0.66	0.744	0.422 +/- 0.021	Facultative	Aquatic
	Shewanella_oneidensis	+	+	-	13.136	0.7	0.695	0.419 +/- 0.021	Facultative	Multiple
	Shigella_boydii_Sb227	+	+	-	8.723	0.61	0.605	0.383 +/- 0.025	Facultative	Host-associated
	Shigella_dysenteriae	+	+	-	8.205	0.57	0.614	0.394 +/- 0.027	Facultative	Host-associated
	Shigella_flexneri_2a	+	+	-	9.735	0.61	0.649	0.4 +/- 0.026	Facultative	Host-associated
	Shigella_flexneri_2a_2457T	+	+	-	10.975	0.63	0.671	0.406 +/- 0.024	Facultative	Host-associated
	Shigella_flexneri_5_8401	+	+	-	9.053	0.67	0.614	0.399 +/- 0.024	Facultative	Host-associated
	Shigella_sonnei_Ss046	+	+	-	11.912	0.6	0.688	0.404 +/- 0.024	Facultative	Host-associated
	Sodalis_glossinidius_morsitans	+	+	-	4.16	0.65	0.556	0.412 +/- 0.035	Microaerophilic	Host-associated
	Thiomicrospira_crunogena_XCL-2	+	-	-	8.329	0.64	0.636	0.372 +/- 0.032	Anaerobic	Aquatic
	Vibrio_cholerae	+	+	-	10.8	0.65	0.673	0.41 +/- 0.024	Facultative	Aquatic
	Vibrio_fischeri_ES114	+	+	-	16.643	0.61	0.789	0.404 +/- 0.023		Multiple
	Vibrio_parahaemolyticus	+	+	-	16.863	0.62	0.76	0.407 +/- 0.021	Facultative	Aquatic
	Vibrio_vulnificus_CMCP6	+	+	-	15.609	0.63	0.755	0.411 +/- 0.022	Facultative	Aquatic
	Vibrio_vulnificus_YJ016	+	+	-	15.081	0.62	0.708	0.411 +/- 0.02	Facultative	Aquatic
	Wigglesworthia_brevipalpis	+	-	-	-1.341	0.77	0.333	0.458 +/- 0.093		Host-associated
	Xanthomonas_campestris	+	+	-	13.483	0.69	0.678	0.415 +/- 0.019	Aerobic	Host-associated
	Xanthomonas_campestris_8004	+	+	-	14.186	0.7	0.697	0.416 +/- 0.02	Aerobic	Host-associated
	Xanthomonas_campestris_vesicatoria_85-10	+	+	-	14.658	0.68	0.69	0.415 +/- 0.019	Aerobic	Host-associated
	Xanthomonas_citri	+	+	-	13.973	0.69	0.687	0.419 +/- 0.019	Aerobic	Host-associated
	Xanthomonas_oryzae_KACC10331	+	+	-	5.464	0.63	0.545	0.425 +/- 0.022	Aerobic	Host-associated
	Xanthomonas_oryzae_MAFF_311018	+	+	-	9.201	0.66	0.62	0.416 +/- 0.022	Aerobic	Host-associated
	Xylella_fastidiosa	+	+	-	5.149	0.6	0.571	0.408 +/- 0.032	Aerobic	Host-associated
	Yersinia_pestis_Antiqua	+	+	-	14.45	0.49	0.697	0.372 +/- 0.022	Facultative	Multiple
	Yersinia_pestis_biovar_Mediaevails	+	+	-	12.515	0.53	0.694	0.386 +/- 0.025	Facultative	Multiple
	Yersinia_pestis_CO92	+	+	-	14.141	0.5	0.713	0.38 +/- 0.024	Facultative	Multiple
	Yersinia_pestis_KIM	+	+	-	12.91	0.53	0.694	0.389 +/- 0.024	Facultative	Multiple
	Yersinia_pestis_Nepal516	+	+	-	13.881	0.52	0.715	0.382 +/- 0.024	Facultative	Multiple
	Yersinia_pseudotuberculosis_IP32953	+	+	-	13.242	0.55	0.696	0.387 +/- 0.023	Facultative	Multiple
	Xylella_fastidiosa_Temecula1	+	+	-	7.772	0.6	0.672	0.412 +/- 0.033	Aerobic	Host-associated
Beta	Azoarcus_sp_EbN1	+	+	+	11.279	0.69	0.679	0.421 +/- 0.023	Facultative	Terrestrial
	Bordetella_bronchiseptica	+	+	-	12.964	0.53	0.634	0.383 +/- 0.019	Aerobic	Host-associated
	Bordetella_parapertussis	+	+	-	11.781	0.55	0.639	0.39 +/- 0.021	Aerobic	Host-associated
	Bordetella_pertussis	+	+	-	11.287	0.56	0.666	0.383 +/- 0.025	Aerobic	Host-associated
	Burkholderia_383	+	+	-	19.076	0.65	0.716	0.403 +/- 0.016	Facultative	Multiple
	Burkholderia_cenocepacia_AU_1054	+	+	-	15.809	0.65	0.676	0.4 +/- 0.017		
	Burkholderia_cenocepacia_HI2424	+	+	-	16.038	0.66	0.675	0.401 +/- 0.017		

Table S1 cont.

	Burkholderia_cepacia_AMMD	+	+	-	18.075	0.65	0.7	0.4 +/- 0.017		
	Burkholderia_mallei_ATCC_23344	+	+	-	7.761	0.65	0.593	0.416 +/- 0.023		Host-associated
	Burkholderia_pseudomallei_1710b	+	+	-	10.732	0.68	0.63	0.422 +/- 0.019	Aerobic	Terrestrial
	Burkholderia_pseudomallei_K96243	+	+	-	14.294	0.65	0.672	0.407 +/- 0.019	Aerobic	Terrestrial
	Burkholderia_thailandensis_E264	+	+	-	12.112	0.68	0.644	0.414 +/- 0.019	Aerobic	Terrestrial
	Burkholderia_xenovorans_LB400	+	+	-	13.454	0.62	0.617	0.398 +/- 0.016	Aerobic	Multiple
	Chromobacterium_violaceum	+	+	-	12.904	0.72	0.691	0.417 +/- 0.021	Facultative	Multiple
	Dechloromonas_aromatica_RCB	+	+	-	14.769	0.8	0.721	0.417 +/- 0.021	Facultative	Multiple
	Methylobacillus_flagellatus_KT	+	+	-	9.824	0.61	0.644	0.373 +/- 0.028	Aerobic	Specialized
	Neisseria_gonorrhoeae_FA_1090	+	+	-	6.522	0.55	0.609	0.38 +/- 0.035	Aerobic	Host-associated
	Neisseria_meningitidis_MC58	+	+	-	8.221	0.53	0.669	0.376 +/- 0.036	Aerobic	Host-associated
	Neisseria_meningitidis_Z2491	+	+	-	7.679	0.56	0.66	0.38 +/- 0.036	Aerobic	Host-associated
	Nitrosomonas_europaea	+	+	-	10.536	0.58	0.721	0.394 +/- 0.031	Aerobic	Multiple
	Nitrosospira_multiformis_ATCC_25196	+	+	-	8.971	0.69	0.656	0.397 +/- 0.029	Aerobic	Terrestrial
	Polaromonas_JS666	+	+	-	13.763	0.59	0.67	0.393 +/- 0.02	Aerobic	Multiple
	Ralstonia_eutropha_H16	+	+	-	14.291	0.57	0.65	0.391 +/- 0.018	Facultative	Specialized
	Ralstonia_eutropha_JMP134	+	+	-	16.372	0.57	0.681	0.389 +/- 0.018	Facultative	Multiple
	Ralstonia_metallidurans_CH34	+	+	-	14.902	0.61	0.656	0.397 +/- 0.017	Facultative	Specialized
	Ralstonia_solanacearum	+	+	-	13.108	0.69	0.664	0.403 +/- 0.02	Aerobic	Multiple
	Rhodoferax_ferrireducens_T118	+	+	-	10.322	0.77	0.649	0.427 +/- 0.022		
	Thiobacillus_denitrificans_ATCC_25259	+	-	+	9.453	0.72	0.642	0.392 +/- 0.026	Facultative	Multiple
Alpha	Agrobacterium_tumefaciens_C58_UWash	+	+	-	12.223	0.77	0.67	0.407 +/- 0.022	Aerobic	Multiple
	Anaplasma_marginale_St_Maries	+	+	-	0.105	0.86	0.479	0.474 +/- 0.049	Aerobic	Host-associated
	Anaplasma_phagocytophilum_HZ	+	+	-	1.431	0.72	0.533	0.461 +/- 0.05	Aerobic	Host-associated
	Bartonella_henselae_Houston-1	+	+	-	6.42	0.62	0.677	0.403 +/- 0.043	Aerobic	Host-associated
	Bartonella_quintana_Toulouse	+	+	-	4.848	0.58	0.628	0.4 +/- 0.047	Aerobic	Host-associated
	Bradyrhizobium_japonicum	+	+	-	15.061	0.81	0.66	0.418 +/- 0.016	Aerobic	Host-associated
	Brucella_abortus_9-941	+	+	-	9.148	0.79	0.67	0.399 +/- 0.03	Facultative	Host-associated
	Brucella_melitensis	+	+	-	7.786	0.85	0.653	0.409 +/- 0.031	Aerobic	Host-associated
	Brucella_melitensis_biovar_Abortus	+	+	-	9.767	0.79	0.687	0.401 +/- 0.029		
	Brucella_suis_1330	+	+	-	8.745	0.82	0.642	0.405 +/- 0.027	Aerobic	Host-associated
	Candidatus_Pelagibacter_ubique_HTCC1062		+	-	7.022	0.78	0.747	0.408 +/- 0.048	Aerobic	Aquatic
	Caulobacter_crescentus	+	+	-	12.223	0.73	0.677	0.405 +/- 0.022	Aerobic	Aquatic
	Ehrlichia_canis_Jake	+	+	-	2.046	0.9	0.58	0.48 +/- 0.049		Host-associated
	Ehrlichia_chaffeensis_Arkansas	+	+	-	1.069	0.77	0.519	0.461 +/- 0.054		Host-associated
	Ehrlichia_ruminantium_Gardel	+	+	-	2.409	0.74	0.603	0.464 +/- 0.058		Host-associated
	Ehrlichia_ruminantium_str._Welgevonden	+	+	-	2.381	0.71	0.587	0.455 +/- 0.055		Host-associated
	Ehrlichia_ruminantium_Welgevonden	+	+	-	2.916	0.72	0.635	0.467 +/- 0.058		Host-associated
	Erythrobacter_litoralis_HTCC2594	+	+	-	15.011	0.8	0.791	0.406 +/- 0.026	Aerobic	Aquatic
	Gluconobacter_oxydans_621H	+	+	-	8.884	0.6	0.674	0.41 +/- 0.03	Aerobic	Multiple
	Granulobacter_bethesdensis_CGDNIH1	+		-	7.794	0.76	0.653	0.41 +/- 0.031		Multiple
	Hyphomonas_neptunium_ATCC_15444	+	+	-	16.445	0.74	0.764	0.404 +/- 0.022	Aerobic	Aquatic
	Jannaschia_CCS1	+	+	-	16.889	0.84	0.797	0.411 +/- 0.023	Aerobic	Aquatic
	Magnetospirillum_magneticum_AMB-1	+	+	-	11.93	0.83	0.686	0.421 +/- 0.022	Microaerophilic	Aquatic
	Maricaulis_maris_MCS10	+	+	-	18.313	0.66	0.809	0.394 +/- 0.023	Facultative	Aquatic
	Mesorhizobium_BNC1	+	+	-	13.866	0.83	0.711	0.411 +/- 0.022	Aerobic	Multiple
	Mesorhizobium_loti	+	+	-	15.067	0.92	0.704	0.417 +/- 0.019	Aerobic	Multiple
	Neorickettsia_sennetsu_Miyayama	+	+	-	0.851	0.83	0.507	0.454 +/- 0.063		Multiple
	Nitrobacter_hamburgensis_X14	+	+	-	13.268	0.89	0.7	0.405 +/- 0.022	Aerobic	Terrestrial
	Nitrobacter_winogradskyi_Nb-255	+	+	-	9.139	0.84	0.652	0.409 +/- 0.027	Facultative	Terrestrial

Table S1 cont.

		Novosphingobium_aromaticivorans_DSM_12444	+	+	-	13.385	0.69	0.701	0.405 +/- 0.022	Aerobic	Multiple
		Rhizobium_etli_CFN_42	+	+	-	13.371	0.79	0.682	0.399 +/- 0.021		Host-associated
		Rhizobium_leguminosarum_bv_viciae_3841	+	+	-	13.117	0.82	0.645	0.41 +/- 0.018	Aerobic	Host-associated
		Rhodobacter_sphaeroides_2_4_1	+	+	-	12.189	0.86	0.701	0.416 +/- 0.023	Facultative	Multiple
		Rhodopseudomonas_palustris_BisA53	+	+	-	15.546	0.79	0.713	0.4 +/- 0.02	Facultative	Multiple
		Rhodopseudomonas_palustris_BisB18	+	+	-	14.2	0.8	0.697	0.408 +/- 0.02	Facultative	Multiple
		Rhodopseudomonas_palustris_BisB5	+	+	-	14.112	0.83	0.718	0.407 +/- 0.022	Facultative	Multiple
		Rhodopseudomonas_palustris_CGA009	+	+	-	13.187	0.73	0.683	0.403 +/- 0.021	Facultative	Multiple
		Rhodopseudomonas_palustris_HaA2	+	+	-	14.47	0.77	0.703	0.403 +/- 0.021	Facultative	Multiple
		Rhodospirillum_rubrum_ATCC_11170	+	+	-	10.344	0.65	0.668	0.399 +/- 0.026	Facultative	Multiple
		Rickettsia_bellii_RML369-C	+	+	-	1.721	0.93	0.505	0.43 +/- 0.044		Host-associated
		Rickettsia_conorii	+	+	-	-0.532	0.69	0.371	0.402 +/- 0.058	Aerobic	Host-associated
		Rickettsia_felis_URRWXCal2	+	+	-	1.831	0.71	0.495	0.402 +/- 0.051		Host-associated
		Rickettsia_prowazekii	+	+	-	0.564	0.7	0.454	0.419 +/- 0.063	Aerobic	Host-associated
		Rickettsia_typhi_wilmington	+	+	-	0.916	0.67	0.476	0.423 +/- 0.058	Aerobic	Host-associated
		Roseobacter_denitrificans_OCh_114	+	+	-	12.323	0.74	0.705	0.409 +/- 0.024		Multiple
		Silicibacter_pomeroyi_DSS-3	+	+	-	12.883	0.78	0.715	0.418 +/- 0.023	Aerobic	Aquatic
		Silicibacter_TM1040	+	+	-	13.815	0.81	0.751	0.415 +/- 0.024		Multiple
		Sinorhizobium_meliloti	+	+	+	14.814	0.78	0.702	0.408 +/- 0.02	Aerobic	Multiple
		Sphingopyxis_alaskensis_RB2256	+	+	+	15.058	0.74	0.762	0.406 +/- 0.024	Aerobic	Aquatic
		Wolbachia_endosymbiont_of_Brugia_malayi_TRS	+	+	-	0.146	0.72	0.465	0.454 +/- 0.075		Host-associated
		Wolbachia_endosymbiont_of_Drosophila_melanogaster	+	+	-	-1.317	0.76	0.363	0.445 +/- 0.062		Host-associated
		Zymomonas_mobilis_ZM4	+	+	-	4.567	0.65	0.571	0.398 +/- 0.038	Anaerobic	Multiple
	Epsilon	Campylobacter_jejuni	+	+	-	7.373	0.56	0.689	0.402 +/- 0.039	Microaerophilic	Multiple
		Campylobacter_jejuni_RM1221	+	+	-	7.137	0.59	0.667	0.404 +/- 0.037	Microaerophilic	Multiple
		Helicobacter_acinonychis_Sheeba	-	+	-	9.32	0.6	0.76	0.406 +/- 0.038	Microaerophilic	Host-associated
		Helicobacter_hepaticus	-	-	-	6.167	0.7	0.642	0.439 +/- 0.033	Aerobic	Host-associated
		Helicobacter_pylori_26695	-	+	-	8.232	0.6	0.708	0.414 +/- 0.036	Aerobic	Host-associated
		Helicobacter_pylori_HPAG1	-	+	-	8.689	0.6	0.725	0.416 +/- 0.036	Aerobic	Host-associated
		Helicobacter_pylori_J99	-	+	-	10.493	0.59	0.777	0.415 +/- 0.035	Aerobic	Host-associated
		Thiomicrospira_denitrificans_ATCC_33889	-		-	4.316	0.82	0.55	0.412 +/- 0.032	Anaerobic	
		Wolinella_succinogenes	-	-	-	4.047	0.83	0.559	0.425 +/- 0.033	Microaerophilic	Host-associated
	Delta	Anaeromyxobacter_dehalogenans_2CP-C	+	-	+	7.427	1.19	0.597	0.461 +/- 0.018	Facultative	Terrestrial
		Bdellovibrio_bacteriovorus	+	-	+	17.105	1.01	0.717	0.444 +/- 0.016	Aerobic	Multiple
		Desulfotalea_psychrophila_LSv54	+	+	-	1.238	0.59	0.462	0.428 +/- 0.028	Anaerobic	Specialized
		Desulfovibrio_desulfuricans_G20	+	-	-	0.991	0.71	0.474	0.448 +/- 0.027	Anaerobic	Multiple
		Desulfovibrio_vulgaris_Hildenborough	+	-	-	1.224	0.66	0.471	0.437 +/- 0.028	Anaerobic	Multiple
		Geobacter_metallireducens_GS-15	+	+	-	3.341	0.9	0.535	0.459 +/- 0.023	Anaerobic	Aquatic
		Geobacter_sulfurreducens	-	-	-	5.073	0.99	0.574	0.457 +/- 0.023	Anaerobic	Multiple
		Lawsonia_intracellularis_PHE_MN1-00	+	-	-	0.395	0.71	0.471	0.452 +/- 0.048	Facultative	Host-associated
		Myxococcus_xanthus_DK_1622	+	-	+	14.853	1.04	0.664	0.458 +/- 0.014	Aerobic	Terrestrial
		Pelobacter_carbinolicus	-	-	-	0.441	0.54	0.438	0.426 +/- 0.028	Anaerobic	Aquatic
		Syntrophobacter_fumaroxidans_MPOB	-	-	-	0.523	0.76	0.468	0.456 +/- 0.024	Anaerobic	Aquatic
		Syntrophus_aciditrophicus_SB	+	-	-	-1.261	0.74	0.401	0.439 +/- 0.03	Anaerobic	Multiple
Acidobacteria	Acidobacteriales	Acidobacteria_bacterium_Ellin345	+	-	+	9.023	0.67	0.571	0.427 +/- 0.016		
	Solibacteres	Solibacter_usitatus_Ellin6076	+	-	-	8.111	0.63	0.519	0.428 +/- 0.011	Aerobic	Terrestrial
Cyanobacteria	Prochlorales	Prochlorococcus_marinus_CCMP1375	-	-	+	7.851	0.86	0.781	0.429 +/- 0.045		Aquatic
		Prochlorococcus_marinus_MED4	-	-	+	6.686	0.83	0.728	0.423 +/- 0.046		Aquatic
		Prochlorococcus_marinus_MIT_9312	-	-	+	6.449	0.8	0.716	0.414 +/- 0.047		Aquatic

Table S1 cont.

		Prochlorococcus_marinus_MIT9313	-	-	+	8.056	0.8	0.703	0.436 +/- 0.033		Aquatic
		Prochlorococcus_marinus_NATL2A	-	-	+	7.273	0.92	0.747	0.408 +/- 0.047		Aquatic
	Chroococcales	Cyanobacteria_bacterium_Yellowstone_A-Prime	+	-	+	5.723	0.62	0.598	0.399 +/- 0.035	Facultative	Specialized
		Cyanobacteria_bacterium_Yellowstone_B-Prime	+	-	+	4.55	0.6	0.55	0.403 +/- 0.032	Facultative	Specialized
		Synechococcus_CC9311	-	-	+	10.151	0.85	0.717	0.423 +/- 0.029		Aquatic
		Synechococcus_CC9605	-	-	+	7.436	0.84	0.664	0.427 +/- 0.032		Aquatic
		Synechococcus_CC9902	-	-	+	8.619	0.83	0.729	0.43 +/- 0.035		Aquatic
		Synechococcus_elongatus_PCC_6301	-	-	+	5.667	0.86	0.614	0.436 +/- 0.031		Aquatic
		Synechococcus_elongatus_PCC_7942	-	-	+	6.756	0.85	0.634	0.43 +/- 0.03		Aquatic
		Synechococcus_sp_WH8102	-	-	+	7.176	0.87	0.673	0.433 +/- 0.033		Aquatic
		Synechocystis_PCC6803	-	-	+	10.459	0.73	0.634	0.413 +/- 0.021		
		Thermosynechococcus_elongatus	-	-	+	4.581	0.69	0.578	0.42 +/- 0.035		Specialized
	Oscillatoriales	Trichodesmium_erythraeum_IMS101	+	-	+	8.1	0.65	0.654	0.4 +/- 0.031	Aerobic	Aquatic
	Nostocales	Anabaena_variabilis_ATCC_29413	+	-	+	11.718	0.58	0.667	0.393 +/- 0.023	Aerobic	Multiple
		Nostoc_sp	+	-	+	11.145	0.59	0.653	0.388 +/- 0.024	Aerobic	Multiple
	Gloeobacteria	Gloeobacter_violaceus	+	-	+	8.518	0.51	0.593	0.398 +/- 0.023		Terrestrial
Deinococcus	Deinococci	Deinococcus_geothermalis_DSM_11300	+	+	-	11.272	0.69	0.676	0.362 +/- 0.028	Aerobic	Aquatic
		Deinococcus_radiodurans	+	+	-	12.028	0.97	0.725	0.404 +/- 0.027	Aerobic	Terrestrial
		Thermus_thermophilus_HB27	+	+	-	9.419	1.21	0.682	0.368 +/- 0.033	Aerobic	Specialized
		Thermus_thermophilus_HB8	+	+	-	11.66	1.16	0.736	0.363 +/- 0.032	Aerobic	Specialized
Aquificae	Aquificales	Aquifex_aeolicus	-	-	-	6.339	0.95	0.696	0.423 +/- 0.043	Aerobic	Specialized
Chloroflexi	Dehalococcoidetes	Dehalococcoides_CBDB1	-	-	-	0.196	1.23	0.487	0.477 +/- 0.053	Anaerobic	Multiple
		Dehalococcoides_ethenogenes_195	-	-	-	-0.094	1.15	0.462	0.467 +/- 0.058	Anaerobic	Multiple
Thermotogae	Thermotogales	Thermotoga_maritima	-	-	-	0.104	0.54	0.375	0.37 +/- 0.044	Anaerobic	Specialized
Fusobacteria	Fusobacterales	Fusobacterium_nucleatum	-	-	-	0.541	0.42	0.338	0.311 +/- 0.051	Anaerobic	Host-associated
Actinobacteria	Actinobacteridae	Arthrobacter_FB24	+	-	+	10.437	0.97	0.702	0.412 +/- 0.028		
		Bifidobacterium_longum	-	-	+	-0.948	0.66	0.371	0.425 +/- 0.056	Anaerobic	Host-associated
		Corynebacterium_diphtheriae	+	+	-	4.861	0.91	0.617	0.423 +/- 0.04	Aerobic	Multiple
		Corynebacterium_efficiens_YS-314	+	-	+	6.12	0.9	0.656	0.408 +/- 0.04	Facultative	Multiple
		Corynebacterium_glutamicum_ATCC_13032_Bielefeld	+	-	+	9.542	1.06	0.731	0.403 +/- 0.034	Facultative	Multiple
		Corynebacterium_glutamicum_ATCC_13032_Kitasato	+	-	+	8.109	1.01	0.683	0.405 +/- 0.034	Facultative	Multiple
		Corynebacterium_jeikeium_K411	+	-	+	7.494	0.9	0.75	0.412 +/- 0.045		
		Frankia_alni_ACN14a	+	-	-	6.289	1.07	0.595	0.427 +/- 0.027		Host-associated
		Frankia_CcI3	+	-	-	5.458	1.07	0.623	0.441 +/- 0.033	Aerobic	Multiple
		Leifsonia_xyli_xyli_CTCB0	+	-	+	4.55	0.99	0.595	0.387 +/- 0.046	Aerobic	Host-associated
		Mycobacterium_avium_paratuberculosis	-	-	+	8.847	1.07	0.673	0.435 +/- 0.027	Aerobic	Multiple
		Mycobacterium_bovis	+	-	+	7.058	1	0.639	0.434 +/- 0.029	Aerobic	Host-associated
		Mycobacterium_leprae	-	-	+	2.986	1.01	0.585	0.446 +/- 0.047	Aerobic	Host-associated
		Mycobacterium_MCS	+	-	+	11.526	1	0.701	0.42 +/- 0.024		
		Mycobacterium_tuberculosis_CDC1551	+	-	+	6.501	0.96	0.615	0.428 +/- 0.029	Aerobic	Host-associated
		Mycobacterium_tuberculosis_H37Rv	+	-	+	6.898	1	0.634	0.435 +/- 0.029	Aerobic	Host-associated
		Nocardia_farcinica_IFM10152	+	-	+	12.135	0.98	0.709	0.415 +/- 0.024	Aerobic	Multiple
		Propionibacterium_acnes_KPA171202	+	+	-	3.006	0.88	0.565	0.447 +/- 0.039	Anaerobic	Host-associated
		Rhodococcus_RHA1	+	-	+	13.935	0.89	0.721	0.421 +/- 0.022	Aerobic	Terrestrial
		Streptomyces_avermitilis	+	-	+	10.214	0.99	0.634	0.441 +/- 0.019	Aerobic	Multiple
		Streptomyces_coelicolor	+	-	+	13.067	0.97	0.681	0.436 +/- 0.019	Aerobic	Multiple
		Thermobifida_fusca_YX	+	-	-	7.468	0.99	0.687	0.433 +/- 0.034	Aerobic	Multiple
		Tropheryma_whipplei_TW08_27	+	-	+	2.062	0.75	0.6	0.455 +/- 0.07	Aerobic	Host-associated
		Tropheryma_whipplei_Twist	+	-	+	1.314	0.71	0.531	0.439 +/- 0.07	Aerobic	Host-associated

Table S1 cont.

	Rubrobacteridae	Rubrobacter_xylanophilus_DSM_9941	+	-	+	4.148	0.71	0.552	0.398 +/- 0.037	Aerobic	Specialized
	Symbiobacterium	Symbiobacterium_thermophilum_IAM14863	+	-	+	5.911	0.67	0.58	0.408 +/- 0.029	Microaerophilic	Terrestrial
Spirochaetes	Spirochaetales	Borrelia_afzelii_PKo	-	-	-	0.499	0.38	0.31	0.283 +/- 0.053	Aerobic	Host-associated
		Borrelia_burgdorferi	-	-	-	2.286	0.44	0.387	0.299 +/- 0.039	Microaerophilic	Host-associated
		Borrelia_garinii_PBi	-	-	-	0.37	0.39	0.296	0.277 +/- 0.05		Host-associated
		Leptospira_borgpetersenii_serovar_Hardjo-bovis_JB197	+	-	+	5.077	0.82	0.55	0.418 +/- 0.026	Aerobic	Host-associated
		Leptospira_borgpetersenii_serovar_Hardjo-bovis_L550	+	-	+	4.941	0.84	0.545	0.42 +/- 0.025	Aerobic	Host-associated
		Leptospira_interrogans_serovar_Copenhageni	+	-	+	9.352	0.94	0.642	0.435 +/- 0.022	Aerobic	Host-associated
		Leptospira_interrogans_serovar_Lai	+	-	+	9.002	0.91	0.628	0.427 +/- 0.022	Aerobic	Host-associated
		Treponema_denticola_ATCC_35405	-	-	-	-1.064	0.71	0.415	0.443 +/- 0.027	Anaerobic	Host-associated
		Treponema_pallidum	-	-	-	0.808	0.65	0.507	0.473 +/- 0.043	Anaerobic	Host-associated
Planctomycetes	Planctomycetacia	Pirellula_sp	-	-	+	0.9	0.76	0.468	0.454 +/- 0.016	Aerobic	Aquatic
Chlamydiae	Chlamydiales	Chlamydia_muridarum	+	+	-	1.427	0.9	0.545	0.476 +/- 0.048		Host-associated
		Chlamydia_trachomatis	+	+	-	-0.961	0.9	0.429	0.476 +/- 0.049		Host-associated
		Chlamydia_trachomatis_A_HAR-13	+	+	-	-0.584	0.91	0.448	0.476 +/- 0.049		Host-associated
		Chlamydophila_abortus_S26_3	+	+	-	-1.22	1.01	0.432	0.489 +/- 0.046		Host-associated
		Chlamydophila_caviae	+	+	-	-0.674	1.01	0.448	0.479 +/- 0.045		Host-associated
		Chlamydophila_felis_Fe_C-56	+	+	-	-1.51	1	0.412	0.48 +/- 0.045		Host-associated
		Chlamydophila_pneumoniae_AR39	+	+	-	-0.065	0.94	0.468	0.471 +/- 0.046		Host-associated
		Chlamydophila_pneumoniae_CWL029	+	+	-	-0.39	0.92	0.459	0.476 +/- 0.043		Host-associated
		Chlamydophila_pneumoniae_J138	+	+	-	-0.294	0.92	0.465	0.478 +/- 0.045		Host-associated
		Chlamydophila_pneumoniae_TW_183	+	+	-	0.002	0.92	0.474	0.474 +/- 0.044		Host-associated
		Parachlamydia_sp_UWE25	-	-	-	-0.678	0.83	0.429	0.453 +/- 0.036		Host-associated
Bacteroidetes	Sphingobacteria	Cytophaga_hutchinsonii_ATCC_33406	+	-	-	0.113	0.96	0.459	0.457 +/- 0.018	Aerobic	Multiple
		Salinibacter_ruber_DSM_13855	+	-	+	7.786	0.38	0.578	0.328 +/- 0.032	Aerobic	Specialized
	Bacteroidetes	Bacteroides_fragilis_NCTC_9434	-	-	-	0.079	0.67	0.462	0.461 +/- 0.015	Anaerobic	Host-associated
		Bacteroides_fragilis_YCH46	-	-	-	-0.43	0.66	0.454	0.46 +/- 0.014	Anaerobic	
		Bacteroides_thetaiotaomicron_VPI-5482	-	-	-	-1.615	0.71	0.448	0.47 +/- 0.013	Anaerobic	Host-associated
		Porphyromonas_gingivalis_W83	-	-	-	-0.627	0.68	0.429	0.448 +/- 0.03	Anaerobic	Host-associated
	Flavobacteriales	Flavobacterium_johnsoniae_UW101	-	-	+	2.512	0.72	0.457	0.419 +/- 0.015	Aerobic	Multiple
Chlorobi	Chlorobia	Chlorobium_chlorochromatii_CaD3	-	-	-	2.297	0.79	0.522	0.435 +/- 0.038	Anaerobic	Aquatic
		Chlorobium_tepidum_TLS	-	-	-	0.638	0.68	0.438	0.415 +/- 0.037	Anaerobic	Specialized
		Pelodictyon_luteolum_DSM_273	-	-	-	-0.477	0.63	0.39	0.407 +/- 0.036	Anaerobic	Multiple
Firmicutes	Lactobacillales	Enterococcus_faecalis_V583	-	-	-	1.859	0.38	0.363	0.287 +/- 0.041	Facultative	Multiple
		Lactobacillus_acidophilus_NCFM	-	-	-	-0.765	0.47	0.237	0.278 +/- 0.054	Facultative	Multiple
		Lactobacillus_brevis_ATCC_367	-	-	-	-0.34	0.45	0.219	0.237 +/- 0.052	Facultative	Multiple
		Lactobacillus_casei_ATCC_334	-	-	-	0.437	0.28	0.194	0.169 +/- 0.056	Facultative	Multiple
		Lactobacillus_delbrueckii_bulgaricus	-	-	-	1.096	0.42	0.355	0.287 +/- 0.062		
		Lactobacillus_delbrueckii_bulgaricus_ATCC_BAA-365	-	-	-	-0.763	0.51	0.265	0.311 +/- 0.06	Facultative	Multiple
		Lactobacillus_gasseri_ATCC_33323	-	-	-	0.534	0.38	0.291	0.251 +/- 0.074	Facultative	Host-associated
		Lactobacillus_johnsonii_NCC_533	-	-	-	1.041	0.39	0.31	0.235 +/- 0.072	Facultative	Host-associated
		Lactobacillus_plantarum	-	-	-	-1.169	0.46	0.213	0.267 +/- 0.046	Facultative	Host-associated
		Lactobacillus_sakei_23K	-	-	-	-0.858	0.48	0.206	0.255 +/- 0.057	Facultative	Multiple
		Lactobacillus_salivarius_UCC118	-	-	-	-0.22	0.49	0.265	0.279 +/- 0.065	Facultative	Host-associated
		Lactococcus_lactis	-	-	-	-0.194	0.41	0.206	0.217 +/- 0.057	Facultative	Multiple
		Lactococcus_lactis_cremoris_SK11	-	-	-	0.647	0.36	0.259	0.222 +/- 0.057	Facultative	Multiple
		Leuconostoc_mesenteroides_ATCC_8293	-	-	-	3.669	0.34	0.462	0.153 +/- 0.084	Facultative	Multiple
		Oenococcus_oeni_PSU-1	-	-	-	-0.003	0.31	0.145	0.145 +/- 0.085	Facultative	Multiple

Table S1 cont.

	Pediococcus_pentosaceus_ATCC_25745	-	-	-	3.223	0.43	0.476	0.235 +/- 0.075	Facultative	Multiple
	Streptococcus_agalactiae_2603	-	-	-	0.826	0.42	0.315	0.269 +/- 0.056	Facultative	Host-associated
	Streptococcus_agalactiae_A909	-	-	-	0.938	0.43	0.324	0.271 +/- 0.056	Facultative	Host-associated
	Streptococcus_agalactiae_NEM316	-	-	-	1.176	0.51	0.359	0.296 +/- 0.054	Facultative	Host-associated
	Streptococcus_mutans	-	-	-	3.306	0.34	0.468	0.252 +/- 0.065	Facultative	Host-associated
	Streptococcus_pneumoniae_D39	-	-	-	1.723	0.41	0.371	0.275 +/- 0.056	Facultative	Multiple
	Streptococcus_pneumoniae_R6	-	-	-	1.514	0.42	0.379	0.294 +/- 0.056	Facultative	Multiple
	Streptococcus_pneumoniae_TIGR4	-	-	-	1.179	0.41	0.351	0.286 +/- 0.055	Facultative	Multiple
	Streptococcus_pyogenes_M1_GAS	-	-	-	1.967	0.52	0.419	0.31 +/- 0.056	Facultative	Host-associated
	Streptococcus_pyogenes_MGAS10270	-	-	-	0.949	0.68	0.401	0.353 +/- 0.051	Facultative	Host-associated
	Streptococcus_pyogenes_MGAS10394	-	-	-	1.422	0.72	0.425	0.354 +/- 0.05	Facultative	Host-associated
	Streptococcus_pyogenes_MGAS10750	-	-	-	1.456	0.72	0.435	0.359 +/- 0.052	Facultative	Host-associated
	Streptococcus_pyogenes_MGAS2096	-	-	-	0.639	0.61	0.371	0.338 +/- 0.052	Facultative	Host-associated
	Streptococcus_pyogenes_MGAS315	-	-	-	1.047	0.64	0.401	0.345 +/- 0.053	Facultative	Host-associated
	Streptococcus_pyogenes_MGAS5005	-	-	-	1.189	0.61	0.394	0.332 +/- 0.052	Facultative	Host-associated
	Streptococcus_pyogenes_MGAS6180	-	-	-	1.487	0.71	0.432	0.36 +/- 0.049	Facultative	Host-associated
	Streptococcus_pyogenes_MGAS8232	-	-	-	1.824	0.62	0.429	0.333 +/- 0.053	Facultative	Host-associated
	Streptococcus_pyogenes_MGAS9429	-	-	-	1.668	0.67	0.432	0.348 +/- 0.05	Facultative	Host-associated
	Streptococcus_pyogenes_SSI-1	-	-	-	2.123	0.65	0.457	0.341 +/- 0.055	Facultative	Host-associated
	Streptococcus_thermophilus_CNRZ1066	-	-	-	2.712	0.37	0.415	0.224 +/- 0.07	Anaerobic	Multiple
	Streptococcus_thermophilus_LMD-9	-	-	-	2.172	0.41	0.401	0.245 +/- 0.072	Facultative	Multiple
	Streptococcus_thermophilus_LMG_18311	-	-	-	1.333	0.32	0.306	0.212 +/- 0.071	Anaerobic	Multiple
Bacillales	Bacillus_anthracis_Ames	+	+	-	4.204	0.34	0.419	0.287 +/- 0.031	Facultative	Multiple
	Bacillus_anthracis_Ames_0581	+	+	-	4.262	0.33	0.419	0.288 +/- 0.031	Facultative	Terrestrial
	Bacillus_anthracis_str_Sterne	+	+	-	3.662	0.35	0.401	0.293 +/- 0.03	Facultative	Multiple
	Bacillus_cereus_ATCC_10987	+	+	-	2.059	0.33	0.355	0.29 +/- 0.031	Aerobic	Terrestrial
	Bacillus_cereus_ATCC14579	+	+	-	3.088	0.34	0.383	0.289 +/- 0.03	Aerobic	Terrestrial
	Bacillus_cereus_ZK	+	+	-	4	0.34	0.405	0.291 +/- 0.029	Aerobic	Terrestrial
	Bacillus_clausii_KSM-K16	-	+	-	4.515	0.39	0.448	0.304 +/- 0.032		
	Bacillus_halodurans	-	+	-	6.352	0.35	0.495	0.281 +/- 0.034	Facultative	Multiple
	Bacillus_licheniformis_ATCC_14580	+	+	-	3.483	0.45	0.432	0.32 +/- 0.032	Facultative	Terrestrial
	Bacillus_licheniformis_DSM_13	+	+	-	3.468	0.43	0.425	0.314 +/- 0.032	Facultative	Terrestrial
	Bacillus_subtilis	+	+	-	2.672	0.44	0.405	0.318 +/- 0.033	Facultative	Terrestrial
	Bacillus_thuringiensis_konkukian	+	+	-	4.042	0.35	0.408	0.293 +/- 0.029	Facultative	Multiple
	Geobacillus_kaustophilus_HTA426	-	+	-	0.623	0.51	0.379	0.356 +/- 0.037	Aerobic	Aquatic
	Listeria_innocua	-	-	-	1.166	0.4	0.32	0.268 +/- 0.045	Facultative	Multiple
	Listeria_monocytogenes	+	-	-	1.634	0.44	0.359	0.282 +/- 0.047	Facultative	Multiple
	Listeria_monocytogenes_4b_F2365	-	-	-	1.475	0.45	0.359	0.29 +/- 0.047	Facultative	Multiple
	Listeria_welshimeri_serovar_6b_SLCC5334	-	-	-	-0.687	0.39	0.231	0.263 +/- 0.047	Facultative	Multiple
	Oceanobacillus_iheyensis	-	+	-	6.423	0.35	0.495	0.246 +/- 0.039	Aerobic	Multiple
	Staphylococcus_aureus_aureus_MRSA252	+	-	-	0.067	0.35	0.248	0.245 +/- 0.047	Facultative	Host-associated
	Staphylococcus_aureus_aureus_MSSA476	+	-	-	-0.566	0.3	0.194	0.22 +/- 0.045	Facultative	Host-associated
	Staphylococcus_aureus_COL	+	-	-	-0.035	0.29	0.225	0.227 +/- 0.05	Facultative	Host-associated
	Staphylococcus_aureus_Mu50	+	-	-	0.672	0.29	0.254	0.224 +/- 0.045	Facultative	Host-associated
	Staphylococcus_aureus_MW2	+	-	-	0.142	0.3	0.225	0.218 +/- 0.047	Facultative	Host-associated
	Staphylococcus_aureus_N315	+	-	-	1.053	0.3	0.275	0.227 +/- 0.046	Facultative	Host-associated
	Staphylococcus_aureus_NCTC_8325	+	-	-	-0.854	0.3	0.187	0.23 +/- 0.05	Facultative	Host-associated
	Staphylococcus_aureus_RF122	+	-	-	0.203	0.39	0.275	0.266 +/- 0.046	Facultative	Host-associated
	Staphylococcus_aureus_USA300	+	-	-	-0.174	0.29	0.206	0.214 +/- 0.045	Facultative	Host-associated
	Staphylococcus_epidermidis_ATCC_12228	+	-	-	-0.873	0.49	0.254	0.298 +/- 0.051	Facultative	Host-associated

Table S1 cont.

		Staphylococcus_epidermidis_RP62A	+	-	-	-0.698	0.42	0.237	0.275 +/- 0.055	Facultative	Host-associated
		Staphylococcus_haemolyticus	+	-	-	-0.827	0.43	0.231	0.277 +/- 0.056	Facultative	Host-associated
		Staphylococcus_saprophyticus	+	-	-	1.343	0.37	0.324	0.248 +/- 0.056	Aerobic	Host-associated
	Mollicutes	Aster_yellows_witches-broom_phytoplasma_AYWB	-	-	-	-1.025	0.94	0.275	0.38 +/- 0.102	Aerobic	Host-associated
		Mesoplasma_florum_L1	-	-	-	-0.954	0.43	0.275	0.358 +/- 0.087	Facultative	Host-associated
		Mycoplasma_capricolum_ATCC_27343	-	-	-	0.778	0.21	0.306	0.251 +/- 0.071	Facultative	Host-associated
		Mycoplasma_gallisepticum	-	-	-	2.788	0.12	0.375	0.173 +/- 0.072	Facultative	Host-associated
		Mycoplasma_genitalium	-	-	-	4.943	0.24	0.875	0.318 +/- 0.113	Facultative	Host-associated
		Mycoplasma_hyopneumoniae_232	-	-	-	3.519	0.22	0.533	0.22 +/- 0.089	Facultative	Host-associated
		Mycoplasma_hyopneumoniae_7448	-	-	-	4.94	0.23	0.643	0.229 +/- 0.084	Facultative	Host-associated
		Mycoplasma_hyopneumoniae_J	-	-	-	4.055	0.25	0.6	0.253 +/- 0.085	Facultative	Host-associated
		Mycoplasma_mobile_163K	-	-	-	1.994	0.13	0.401	0.14 +/- 0.131	Facultative	Host-associated
		Mycoplasma_mycoides	-	-	-	0.025	0.23	0.259	0.257 +/- 0.077	Facultative	Host-associated
		Mycoplasma_penetrans	-	-	-	-0.679	0.29	0.291	0.33 +/- 0.057	Facultative	Host-associated
		Mycoplasma_pneumoniae	-	-	-	0.699	0.24	0.333	0.28 +/- 0.076	Facultative	Host-associated
		Mycoplasma_pulmonis	-	-	-	3.702	0.13	0.454	0.137 +/- 0.086	Facultative	Host-associated
		Mycoplasma_synoviae_53	-	-	-	4.73	0.1	0.667	0.12 +/- 0.116	Facultative	Host-associated
		Onion_yellows_phytoplasma	-	-	-	0.839	0.61	0.454	0.374 +/- 0.096	Aerobic	Host-associated
		Ureaplasma_urealyticum	-	-	-	2.381	0.29	0.535	0.351 +/- 0.077	Facultative	Host-associated
	Clostridia	Carboxydothermus_hydrogenoformans_Z-2901	-	-	-	-0.68	1.14	0.422	0.449 +/- 0.04	Anaerobic	Aquatic
		Clostridium_acetobutylicum	-	-	-	-0.264	0.64	0.432	0.44 +/- 0.031	Anaerobic	Multiple
		Clostridium_perfringens	+	-	-	-1.458	0.47	0.342	0.401 +/- 0.04	Anaerobic	Multiple
		Clostridium_perfringens_ATCC_13124	+	-	-	-1.18	0.48	0.359	0.405 +/- 0.039	Anaerobic	Multiple
		Clostridium_perfringens_SM101	-	-	-	-1.038	0.53	0.363	0.41 +/- 0.045	Anaerobic	Multiple
		Clostridium_tetani_E88	-	-	-	0.939	0.53	0.457	0.42 +/- 0.039	Anaerobic	Multiple
		Desulfitobacterium_hafniense_Y51	-	-	-	-2.091	0.69	0.387	0.441 +/- 0.026	Anaerobic	Specialized
		Moorella_thermoacetica_ATCC_39073	-	-	-	0.257	0.63	0.438	0.428 +/- 0.04	Anaerobic	Aquatic
		Syntrophomonas_wolfei_Goettingen	-	-	-	2.622	0.72	0.548	0.451 +/- 0.037	Anaerobic	Multiple
		Thermoanaerobacter_tengcongensis	-	-	-	0.884	0.57	0.412	0.373 +/- 0.044	Anaerobic	Specialized
Nanoarchaeota	Nanoarchaeum	Nanoarchaeum_equitans	+	-	-	4.292	2.56	0.885	0.488 +/- 0.092	Anaerobic	Host-associated
Crenarchaeota	Thermoprotei	Aeropyrum_pernix	-	-	+	1.796	1.12	0.468	0.395 +/- 0.041	Aerobic	Specialized
		Pyrobaculum_aerophilum	-	-	+	5.986	1.12	0.691	0.437 +/- 0.042	Facultative	Aquatic
		Sulfolobus_acidocaldarius_DSM_639	-	-	+	0.145	0.82	0.371	0.358 +/- 0.09	Aerobic	Specialized
		Sulfolobus_solfataricus	-	-	+	1.245	0.95	0.468	0.389 +/- 0.064	Aerobic	Specialized
		Sulfolobus_tokodaii	-	-	-	0.081	1	0.401	0.396 +/- 0.06	Aerobic	Specialized
Euryarchaeota	Archaeoglobi	Archaeoglobus_fulgidus	+	-	-	-0.17	1.22	0.457	0.465 +/- 0.045	Anaerobic	Aquatic
	Halobacteria	Haloarcula_marismortui_ATCC_43049	+	+	-	4.832	0.58	0.554	0.356 +/- 0.041	Aerobic	Aquatic
		Halobacterium_sp	-	-	-	1.394	0.77	0.454	0.381 +/- 0.052	Facultative	Specialized
		Haloquadratum_walsbyi	-	-	-	3.894	0.4	0.563	0.306 +/- 0.066		
		Natronomonas_pharaonis	+	+	+	3.904	0.47	0.535	0.326 +/- 0.054	Aerobic	Aquatic
	Thermoplasmata	Picrophilus_torridus_DSM_9790	-	-	-	-0.95	1.01	0.342	0.424 +/- 0.086	Aerobic	Specialized
		Thermoplasma_acidophilum	-	-	-	1.122	0.75	0.479	0.388 +/- 0.081	Facultative	Specialized
		Thermoplasma_volcanium	-	-	-	-0.836	0.7	0.275	0.352 +/- 0.092	Facultative	Specialized
	Methanobacteria	Methanobacterium_thermoautotrophicum	-	-	-	1.827	0.75	0.531	0.43 +/- 0.055	Anaerobic	Specialized
		Methanosphaera_stadtmanae	-	-	-	-0.497	0.5	0.419	0.45 +/- 0.061	Anaerobic	Host-associated
	Methanococci	Methanococcus_jannaschii	-	-	-	0.462	0.81	0.476	0.448 +/- 0.061	Anaerobic	Aquatic
		Methanococcus_maripaludis_S2	-	-	-	-0.824	0.97	0.425	0.467 +/- 0.05	Anaerobic	Aquatic
	Methanomicrobia	Methanococcoides_burtonii_DSM_6242	-	-	-	1.359	0.61	0.487	0.414 +/- 0.054	Anaerobic	Aquatic
		Methanosaeta_thermophila_PT	-	-	-	-1.386	0.78	0.387	0.447 +/- 0.043	Anaerobic	
		Methanosarcina_acetivorans	-	-	-	-0.092	0.71	0.441	0.444 +/- 0.03	Anaerobic	Aquatic

Table S1 cont.

		Methanosarcina_barkeri_fusaro	-	-	-	1.099	0.83	0.495	0.461 +/- 0.031		
		Methanosarcina_mazei	-	-	-	0.443	0.77	0.465	0.446 +/- 0.043	Anaerobic	Multiple
		Methanospirillum_hungatei_JF-1	-	-	-	1.451	0.64	0.49	0.445 +/- 0.031	Anaerobic	Multiple
	Methanopyri	Methanopyrus_kandleri	-	-	-	0.325	0.98	0.49	0.473 +/- 0.052	Anaerobic	Specialized
	Thermococci	Pyrococcus_abyssi	-	-	-	-0.06	0.99	0.398	0.402 +/- 0.063	Anaerobic	Aquatic
		Pyrococcus_furiosus	-	-	-	0.929	0.88	0.438	0.392 +/- 0.05	Anaerobic	Aquatic
		Pyrococcus_horikoshii	-	-	-	2.283	1.03	0.485	0.377 +/- 0.047	Anaerobic	Aquatic
		Thermococcus_kodakaraensis_KOD1	-	-	-	2.086	1	0.495	0.4 +/- 0.045	Anaerobic	Specialized

Table S1 cont.