Supplementary Information 1

Metabolic Interactions of a Chain Elongation Microbiome

Wenhao Han[a,b], Pinjing He[b,c,d], Liming Shao[b,c,d], Fan Lü[a,c,*]

[a] *State Key Laboratory of Pollution Control and Resource Reuse, Tongji University, Shanghai 200092, China*

[b] *Shanghai Institute of Pollution Control and Ecological Security, Shanghai 200092, P.R. China*

[c] *Institute of Waste Treatment and Reclamation, Tongji University, Shanghai 200092, China*

[d] *Centre for the Technology Research and Training on Household Waste in Small Towns & Rural Area, Ministry of Housing and Urban-Rural Development of PR China (MOHURD), China*

[*] *Corresponding author: Fan Lü (lvfan.rhodea@tongji.edu.cn)*

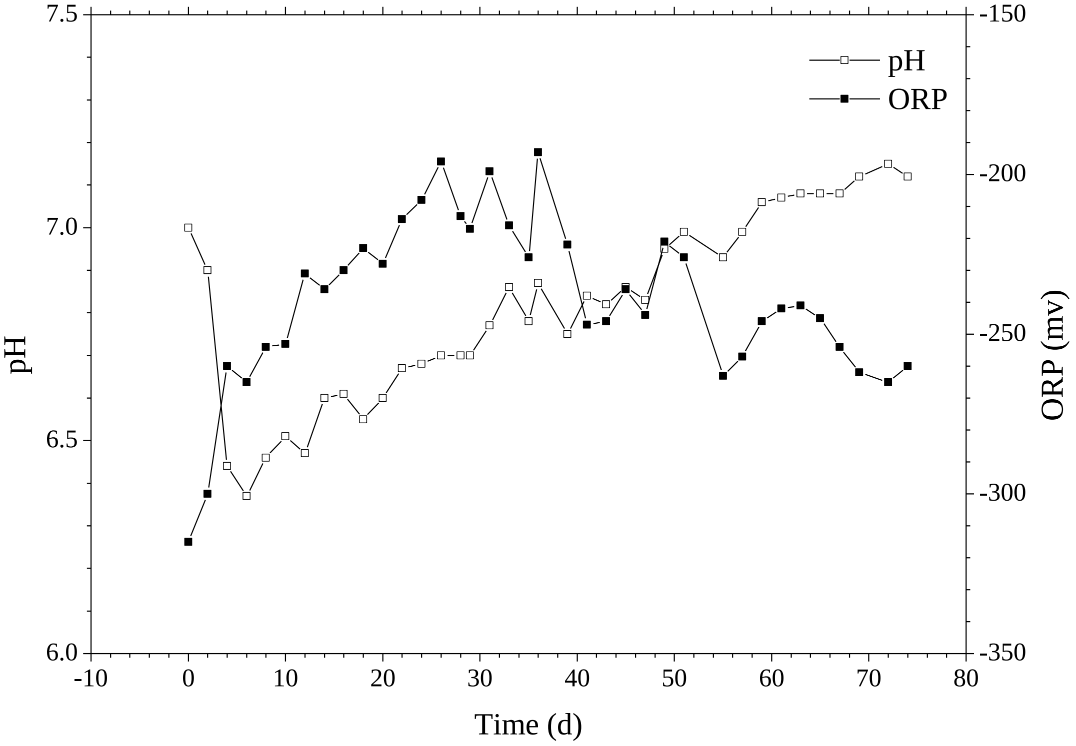

Figure S1. Changes of pH and oxidation-reduction potential (ORP) during the operation period.

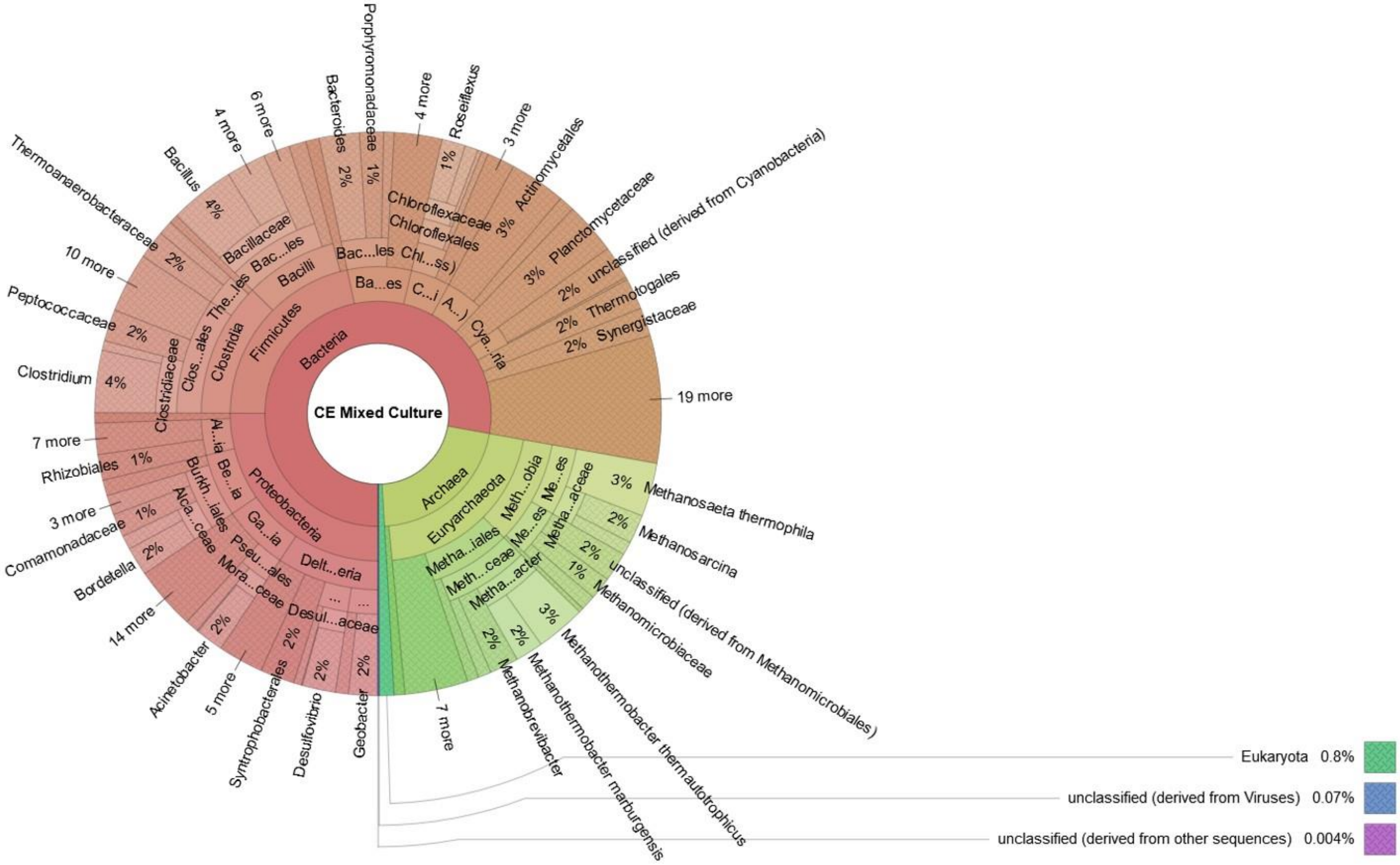

Figure S2. Whole interactive krona chart of the full taxonomy based on DNA sequences.

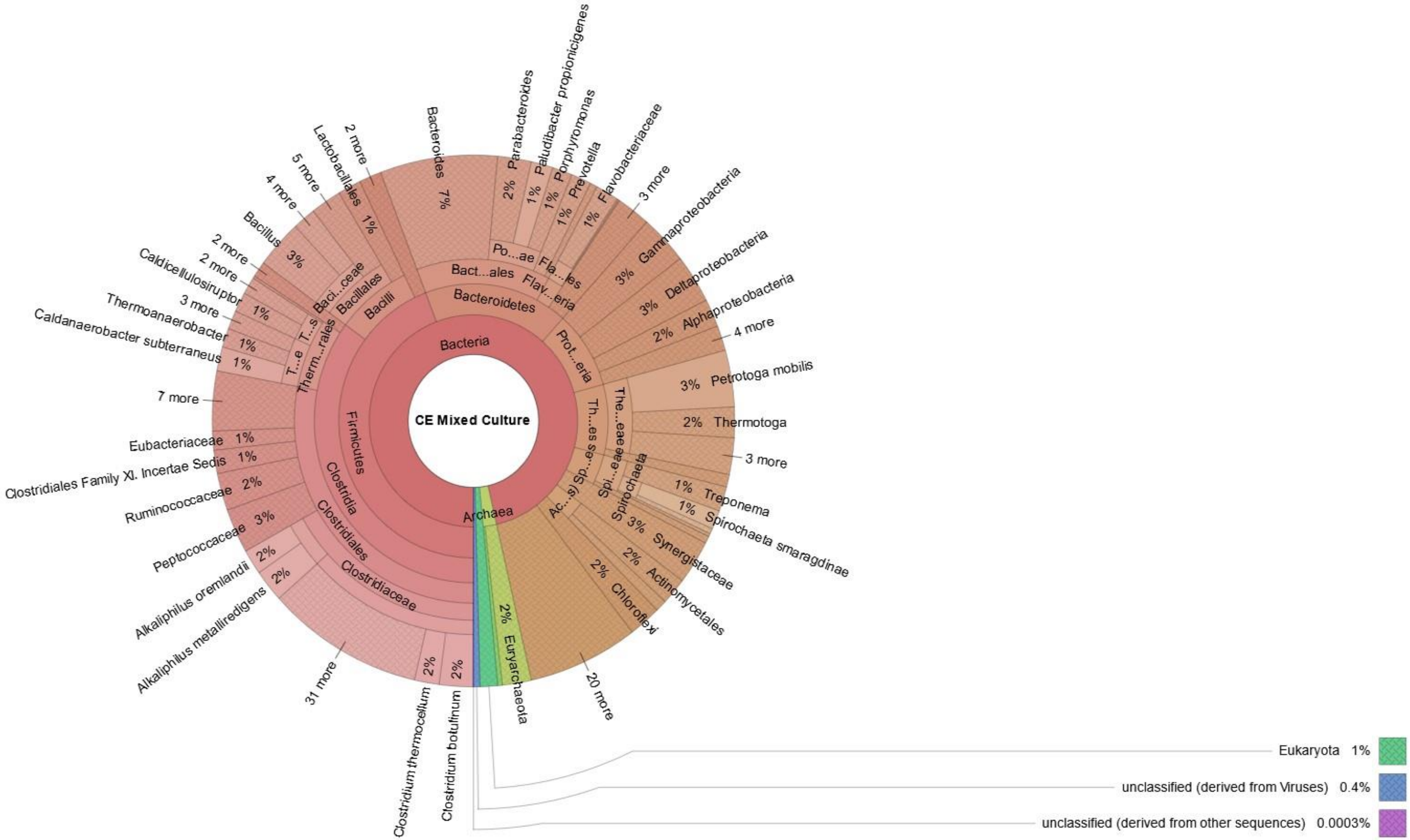

Figure S3. Whole interactive krona chart of the full taxonomy based on cDNA sequence.

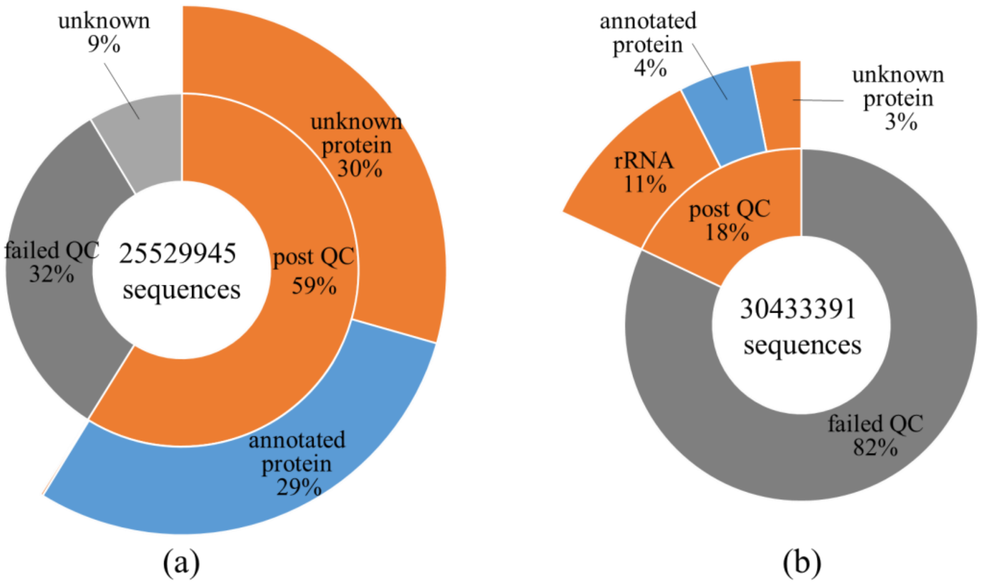

Figure S4. DNA and cDNA sequence quality. (a) DNA, (b) cDNA; average: 150 bp/sequence.

Table S1. Gas production.

Time (day)	4	8	16	20	24	Total
Gas volume (L)	1.882	0.623	0.324	0.154	0	2.983
CH_4 concentration (%)	5.68	7.39	0.90	0.64	0.00	/
CO_2 concentration (%)	1.04	0.96	2.35	0.25	0.00	/
H_2 concentration (%)	7.36	37.23	0.00	0.00	0.00	/
CH_4 production (mmol)	4.23	1.82	0.12	0.04	0.00	6.21
CO_2 production (mmol)	0.77	0.23	0.30	0.02	0.00	1.33
H_2 production (mmol)	5.48	9.18	0.00	0.00	0.00	14.66
CH_4 production / substrate* (%, Base on carbon)	0.15	0.06	0.00	0.00	0.00	0.22
CO_2 production / substrate* (%, Base on carbon)	0.03	0.01	0.01	0.00	0.00	0.05

*substrate: 60 mM ethanol + 20 mM acetate

Table S2. Annotated gene hits against different databases available at MG-RAST.

	Database	Metagenome	Metatranscriptome
Protein	RefSeq	7,688,887	1,254,074
	IMG	6,619,183	1,089,161
	TrEMBL	6,522,003	1,103,386
	SEED	4,560,118	676,086
	KEEG	6,466,532	937,965
	GenBank	7,020,441	1,184,626
	SwissProt	/	178,761
	PATRIC	/	975,387
Hierarchical	Subststems	3,227,567	534,629
	KO	2,931,825	499,106
	COG	2,885,237	404,565
	NOG	28,584	6,014
rRNA	RDP	10,182	2,542,385
	Silva LSU	/	2,312,670
	Silva SSU	/	2,801,991
	Greengenes	8,185	1,691,280

Table S3. Taxonomic affiliation at different levels based on DNA sequences from RefSeq database.

Domain	Bacteria	78.1%
	Archaea	20.9%
	Eukaryota	0.9%
	Viruses	0.1%
Phylum	Proteobacteria	25.3%
	Firmicutes	21.6%
	Euryarchaeota	20.2%
	Bacteroidetes	6.9%
	Actinobacteria	4.3%
	Chloroflexi	4.2%
	Planctomycetes	2.9%
	Cyanobacteria	2.2%
	Thermotogae	1.8%
	Synergistetes	1.6%
	Acidobacteria	0.9%
	Chlorobi	0.9%
	Verrucomicrobia	0.9%
	Spirochaetes	0.9%
	Deinococcus-Thermus	0.8%
	Crenarchaeota	0.5%
	Aquificae	0.5%
	unclassified (derived from Bacteria)	0.5%
	Fusobacteria	0.4%
	Nitrospirae	0.4%
Class	Clostridia	12.3%
	Methanomicrobia	9.3%
	Deltaproteobacteria	9.2%
	Bacilli	8.5%
	Methanobacteria	7.2%
	Gammaproteobacteria	6.2%
	Betaproteobacteria	5.8%
	Actinobacteria (class)	4.3%
	Bacteroidia	4.2%
	Alphaproteobacteria	3.5%
	Planctomycetacia	2.9%
	Chloroflexi (class)	2.7%
	unclassified (derived from Cyanobacteria)	2.0%
	Thermotogae (class)	1.8%

	Synergistia	1.6%
	Methanococci	1.2%
	Flavobacteria	1.1%
	Chlorobia	0.9%
	Spirochaetes (class)	0.9%
	Sphingobacteria	0.8%
Order	Clostridiales	9.3%
	Bacillales	7.5%
	Methanobacteriales	7.2%
	Methanosarcinales	5.4%
	Burkholderiales	4.9%
	Bacteroidales	4.2%
	Methanomicrobiales	3.7%
	Actinomycetales	3.4%
	Planctomycetales	2.9%
	Thermoanaerobacterales	2.4%
	Pseudomonadales	2.4%
	Desulfuromonadales	2.4%
	Desulfovibrionales	2.3%
	Chloroflexales	2.3%
	Syntrophobacterales	1.9%
	Thermotogales	1.8%
	Synergistales	1.6%
	Rhizobiales	1.6%
	Methanococcales	1.2%
	Chroococcales	1.1%
Family	Methanobacteriaceae	6.6%
	Bacillaceae	5.9%
	Clostridiaceae	4.1%
	Methanosaetaceae	3.0%
	Planctomycetaceae	2.9%
	Methanosarcinaceae	2.4%
	Alcaligenaceae	2.4%
	Bacteroidaceae	2.3%
	Chloroflexaceae	2.1%
	Desulfovibrionaceae	1.9%
	Peptococcaceae	1.8%
	Geobacteraceae	1.6%
	Moraxellaceae	1.6%
	unclassified (derived from Methanomicrobiales)	1.6%
	Thermoanaerobacteraceae	1.6%
	Synergistaceae	1.6%

	Porphyromonadaceae	1.3%
	Comamonadaceae	1.3%
	Thermotogaceae	1.3%
	Methanomicrobiaceae	1.3%
	unclassified (derived from Chroococcales)	1.1%
	Flavobacteriaceae	1.0%
	Syntrophobacteraceae	1.0%
	Syntrophaceae	0.9%
	Burkholderiaceae	0.9%
	Chlorobiaceae	0.9%
	Enterobacteriaceae	0.8%
	Ruminococcaceae	0.8%
	Pseudomonadaceae	0.8%
	Desulfobacteraceae	0.7%
Genus	*Methanothermobacter*	4.3%
	Bacillus	3.7%
	Clostridium	3.6%
	Methanosaeta	3.0%
	Bacteroides	2.3%
	Desulfovibrio	1.9%
	Geobacter	1.6%
	Methanosarcina	1.6%
	Methanobrevibacter	1.6%
	Bordetella	1.6%
	Acinetobacter	1.6%
	Roseiflexus	1.3%
	Lysinibacillus	1.0%
	Syntrophobacter	1.0%
	Syntrophus	0.9%
	Methanoregula	0.9%
	Methanoculleus	0.9%
	Geobacillus	0.8%
	Achromobacter	0.8%
	Chloroflexus	0.7%
	Pirellula	0.7%
	Methanosphaera	0.7%
	Methanosphaerula	0.7%
	Methanocaldococcus	0.7%
	Blastopirellula	0.7%
	Pseudomonas	0.7%
	Planctomyces	0.7%
	Parabacteroides	0.6%

	Methanothermus	0.6%
	Methanospirillum	0.6%
Species	*Methanosaeta thermophila*	3.0%
	Methanothermobacter thermautotrophicus	2.7%
	Methanothermobacter marburgensis	1.5%
	Methanobrevibacter smithii	1.0%
	Syntrophobacter fumaroxidans	1.0%
	Syntrophus aciditrophicus	0.9%
	Methanoregula boonei	0.9%
	Methanoculleus marisnigri	0.9%
	Pirellula staleyi	0.7%
	Methanosphaera stadtmanae	0.7%
	Lysinibacillus sphaericus	0.7%
	Methanosphaerula palustris	0.7%
	Roseiflexus sp. RS-1	0.7%
	Blastopirellula marina	0.7%
	Methanobrevibacter ruminantium	0.6%
	Methanothermus fervidus	0.6%
	Methanospirillum hungatei	0.6%
	Rhodopirellula baltica	0.6%
	Roseiflexus castenholzii	0.6%
	Bacillus sp. B14905	0.6%
	Methanosarcina acetivorans	0.6%
	Bordetella avium	0.6%
	Methanosarcina barkeri	0.6%
	Acinetobacter baumannii	0.6%
	Desulfovibrio desulfuricans	0.5%
	Desulfitobacterium hafniense	0.5%
	Thermotogales bacterium mesG1.Ag.4.2	0.5%
	Methanosarcina mazei	0.5%
	Bordetella bronchiseptica	0.5%
	Candidatus Solibacter usitatus	0.5%
Strain	*Methanosaeta thermophila* PT	3.0%
	Methanothermobacter thermautotrophicus str. Delta H	2.7%
	Methanothermobacter marburgensis str. Marburg	1.5%
	Syntrophobacter fumaroxidans MPOB	1.0%
	Syntrophus aciditrophicus SB	0.9%
	Candidatus Methanoregula boonei 6A8	0.9%
	Methanoculleus marisnigri JR1	0.9%
	Methanobrevibacter smithii ATCC 35061	0.8%
	Pirellula staleyi DSM 6068	0.7%

Methanosphaera stadtmanae DSM 3091	0.7%
Lysinibacillus sphaericus C3-41	0.7%
Methanosphaerula palustris E1-9c	0.7%
Roseiflexus sp. RS-1	0.7%
Blastopirellula marina DSM 3645	0.7%
Methanobrevibacter ruminantium M1	0.6%
Methanothermus fervidus DSM 2088	0.6%
Methanospirillum hungatei JF-1	0.6%
Rhodopirellula baltica SH 1	0.6%
Roseiflexus castenholzii DSM 13941	0.6%
Bacillus sp. B14905	0.6%
Methanosarcina acetivorans C2A	0.6%
Bordetella avium 197N	0.6%
Methanosarcina barkeri str. Fusaro	0.6%
Thermotogales bacterium mesG1.Ag.4.2	0.5%
Methanosarcina mazei Go1	0.5%
Bordetella bronchiseptica RB50	0.5%
Candidatus Solibacter usitatus Ellin6076	0.5%
Achromobacter xylosoxidans A8	0.4%
Methanocaldococcus jannaschii DSM 2661	0.4%
uncultured *methanogenic archaeon* RC-I	0.4%

Note: only top 20 at phylum, class and order level; top 30 at family, genus, species and strain level are listed.

Table S4. Taxonomic affiliation at different levels based on cDNA sequences from RefSeq database.

Domain	Bacteria	96.3%
	Archaea	2.1%
	Eukaryota	1.1%
	Viruses	0.4%
Phylum	Firmicutes	43.9%
	Bacteroidetes	17.1%
	Proteobacteria	9.7%
	Thermotogae	8.0%
	Spirochaetes	3.5%
	Synergistetes	2.6%
	Actinobacteria	2.5%
	Chloroflexi	2.0%
	Euryarchaeota	1.8%
	Fusobacteria	1.1%
	Cyanobacteria	1.0%
	Lentisphaerae	0.7%
	Chlorobi	0.6%
	Verrucomicrobia	0.5%
	Planctomycetes	0.5%
	Deinococcus-Thermus	0.4%
	unclassified (derived from Viruses)	0.4%
	Acidobacteria	0.4%
	unclassified (derived from Bacteria)	0.4%
	Deferribacteres	0.3%
Class	Clostridia	34.6%
	Bacteroidia	13.4%
	Thermotogae (class)	8.0%
	Bacilli	7.9%
	Spirochaetes (class)	3.5%
	Gammaproteobacteria	3.2%
	Deltaproteobacteria	3.0%
	Synergistia	2.6%
	Actinobacteria (class)	2.5%
	Alphaproteobacteria	1.8%
	Flavobacteria	1.6%
	Betaproteobacteria	1.1%
	Fusobacteria (class)	1.1%
	Chloroflexi (class)	1.0%

	Sphingobacteria	1.0%
	unclassified (derived from Cyanobacteria)	0.9%
	Negativicutes	0.9%
	Cytophagia	0.9%
	Dehalococcoidetes	0.7%
	unclassified (derived from Lentisphaerae)	0.7%
Order	Clostridiales	27.6%
	Bacteroidales	13.4%
	Thermotogales	8.0%
	Bacillales	6.4%
	Thermoanaerobacterales	6.0%
	Spirochaetales	3.5%
	Synergistales	2.6%
	Actinomycetales	1.8%
	Flavobacteriales	1.6%
	Lactobacillales	1.5%
	Fusobacteriales	1.1%
	Sphingobacteriales	1.0%
	Desulfuromonadales	0.9%
	Chloroflexales	0.9%
	Selenomonadales	0.9%
	Cytophagales	0.9%
	Desulfovibrionales	0.8%
	Burkholderiales	0.7%
	Enterobacteriales	0.7%
	Rhizobiales	0.7%
Family	Clostridiaceae	16.5%
	Thermotogaceae	7.2%
	Bacteroidaceae	7.1%
	Porphyromonadaceae	4.5%
	Bacillaceae	4.4%
	Thermoanaerobacteraceae	3.7%
	Spirochaetaceae	3.1%
	Peptococcaceae	2.7%
	Synergistaceae	2.6%
	Ruminococcaceae	2.2%
	Thermoanaerobacterales Family III. Incertae Sedis	2.2%
	Flavobacteriaceae	1.5%
	Clostridiales Family XI. Incertae Sedis	1.4%
	Prevotellaceae	1.3%
	Eubacteriaceae	1.2%
	Lachnospiraceae	1.1%

	Fusobacteriaceae	1.1%
	Paenibacillaceae	1.0%
	Syntrophomonadaceae	1.0%
	Chloroflexaceae	0.8%
	unclassified (derived from Thermotogales)	0.7%
	Enterobacteriaceae	0.7%
	Cytophagaceae	0.7%
	Veillonellaceae	0.7%
	unclassified (derived from Dehalococcoidetes)	0.7%
	Desulfovibrionaceae	0.7%
	Chlorobiaceae	0.6%
	Victivallaceae	0.6%
	Sphingobacteriaceae	0.6%
	Halanaerobiaceae	0.5%
Genus	*Clostridium*	13.2%
	Bacteroides	7.1%
	Alkaliphilus	3.4%
	Petrotoga	3.3%
	Bacillus	3.0%
	Parabacteroides	2.0%
	Thermotoga	1.8%
	Treponema	1.5%
	Caldanaerobacter	1.4%
	Spirochaeta	1.4%
	Paludibacter	1.3%
	Prevotella	1.3%
	Porphyromonas	1.2%
	Caldicellulosiruptor	1.2%
	Thermoanaerobacter	1.1%
	Aminobacterium	1.0%
	Eubacterium	1.0%
	Desulfitobacterium	1.0%
	Thermosipho	1.0%
	Geobacillus	0.9%
	Desulfotomaculum	0.9%
	Paenibacillus	0.8%
	unclassified (derived from Thermotogales)	0.7%
	Fervidobacterium	0.7%
	Desulfovibrio	0.7%
	Ruminococcus	0.6%
	Victivallis	0.6%
	Dehalococcoides	0.6%

	Fusobacterium	0.6%
	Syntrophomonas	0.6%
Species	*Petrotoga mobilis*	3.3%
	Clostridium botulinum	2.1%
	Alkaliphilus metalliredigens	1.9%
	Clostridium thermocellum	1.5%
	Alkaliphilus oremlandii	1.5%
	Caldanaerobacter subterraneus	1.4%
	Paludibacter propionicigenes	1.3%
	Spirochaeta smaragdinae	1.1%
	Aminobacterium colombiense	1.0%
	Desulfitobacterium hafniense	1.0%
	Clostridium difficile	0.9%
	Porphyromonas gingivalis	0.9%
	Bacteroides fragilis	0.9%
	Bacteroides sp. 20_3	0.9%
	Thermotogales bacterium mesG1.Ag.4.2	0.7%
	Caldicellulosiruptor saccharolyticus	0.7%
	Thermosipho melanesiensis	0.7%
	Fervidobacterium nodosum	0.7%
	Clostridium kluyveri	0.7%
	Treponema denticola	0.7%
	Clostridium carboxidivorans	0.6%
	Bacteroides sp. 1_1_6	0.6%
	Victivallis vadensis	0.6%
	Desulfotomaculum reducens	0.6%
	Parabacteroides distasonis	0.6%
	Clostridium phytofermentans	0.6%
	Clostridium cellulolyticum	0.6%
	Thermotoga maritima	0.6%
	Clostridium sticklandii	0.6%
	Clostridium perfringens	0.6%
Strain	*Petrotoga mobilis* SJ95	3.3%
	Alkaliphilus metalliredigens QYMF	1.9%
	Alkaliphilus oremlandii OhILAs	1.5%
	Clostridium thermocellum ATCC 27405	1.4%
	Thermoanaerobacter tengcongensis MB4	1.3%
	Paludibacter propionicigenes WB4	1.3%
	Spirochaeta smaragdinae DSM 11293	1.1%
	Aminobacterium colombiense DSM 12261	1.0%
	Bacteroides sp. 20_3	0.9%
	Thermotogales bacterium mesG1.Ag.4.2	0.7%

Desulfitobacterium hafniense Y51	0.7%
Caldicellulosiruptor saccharolyticus DSM 8903	0.7%
Thermosipho melanesiensis BI429	0.7%
Fervidobacterium nodosum Rt17-B1	0.7%
Treponema denticola ATCC 35405	0.7%
Clostridium carboxidivorans P7	0.6%
Bacteroides sp. 1_1_6	0.6%
Victivallis vadensis ATCC BAA-548	0.6%
Desulfotomaculum reducens MI-1	0.6%
Parabacteroides distasonis ATCC 8503	0.6%
Clostridium phytofermentans ISDg	0.6%
Clostridium cellulolyticum H10	0.6%
Thermotoga maritima MSB8	0.6%
Clostridium sticklandii DSM 519	0.6%
Syntrophomonas wolfei subsp. wolfei str. Goettingen	0.6%
Clostridium botulinum A str. ATCC 3502	0.6%
Clostridium kluyveri DSM 555	0.6%
Dethiosulfovibrio peptidovorans DSM 11002	0.6%
Thermoanaerobacter sp. X513	0.5%
Clostridium acetobutylicum ATCC 824	0.5%

Note: only top 20 at phylum, class and order level; top 30 at family, genus, species and strain level are listed.

Table S5. Categories of functional genes at different levels based on DNA sequences from Subsystems database (only top 30 at Level 2, top 50 at Level 3 and Function were listed, Level 1 see Figure 5)

Level 2	NULL	19.8%
	Protein biosynthesis	7.3%
	Plant-Prokaryote DOE project	5.4%
	Central carbohydrate metabolism	4.0%
	RNA processing and modification	3.4%
	DNA repair	2.8%
	Lysine, threonine, methionine, and cysteine	2.8%
	Electron donating reactions	2.3%
	One-carbon Metabolism	2.1%
	Purines	2.0%
	Folate and pterines	2.0%
	Branched-chain amino acids	1.8%
	Aromatic amino acids and derivatives	1.8%
	Resistance to antibiotics and toxic compounds	1.8%
	DNA replication	1.7%
	ABC transporters	1.6%
	Tetrapyrroles	1.5%
	Protein degradation	1.5%
	Capsular and extracellular polysacchrides	1.5%
	Transcription	1.5%
	Monosaccharides	1.5%
	Arginine; urea cycle, polyamines	1.5%
	Di- and oligosaccharides	1.0%
	Pyrimidines	1.0%
	Fermentation	0.9%
	Protein folding	0.9%
	Protein processing and modification	0.9%
	Oxidative stress	0.9%
	Glutamine, glutamate, aspartate, asparagine; ammonia assimilation	0.8%
	Gram-Negative cell wall components	0.8%
Level 3	Methanogenesis	1.0%
	Methionine_Biosynthesis	0.9%
	De_Novo_Purine_Biosynthesis	0.9%
	Serine-glyoxylate_cycle	0.9%
	DNA_repair,_UvrABC_system	0.9%

Purine_conversions	0.9%
tRNA_modification_Archaea	0.9%
Coenzyme_B12_biosynthesis	0.9%
Branched-Chain_Amino_Acid_Biosynthesis	0.8%
Sugar_utilization_in_Thermotogales	0.8%
DNA_repair,_bacterial	0.7%
Arginine_Biosynthesis_extended	0.7%
Ribosome_LSU_bacterial	0.7%
DNA-replication	0.7%
Phosphate_metabolism	0.7%
YgfZ	0.7%
RNA_polymerase_bacterial	0.7%
Hydrogenases	0.6%
ABC_transporter_oligopeptide_(TC_3.A.1.5.1)	0.6%
Glycolysis_and_Gluconeogenesis	0.6%
Pyruvate_metabolism_I:_anaplerotic_reactions,_PEP	0.6%
Proteolysis_in_bacteria,_ATP-dependent	0.6%
Glutamine,_Glutamate,_Aspartate_and_Asparagine_Biosynthesis	0.6%
Universal_GTPases	0.6%
Bacterial_Cell_Division	0.6%
Methanogenesis_strays	0.6%
Peptidoglycan_Biosynthesis	0.6%
Ammonia_assimilation	0.6%
Respiratory_Complex_I	0.6%
Heat_shock_dnaK_gene_cluster_extended	0.5%
DNA_replication,_archaeal	0.5%
Histidine_Biosynthesis	0.5%
PROSC	0.5%
Protein_chaperones	0.5%
De_Novo_Pyrimidine_Synthesis	0.5%
Ribosome_SSU_bacterial	0.5%
Cell_division-ribosomal_stress_proteins_cluster	0.5%
Restriction-Modification_System	0.5%
Chorismate:_Intermediate_for_synthesis_of_Tryptophan,_PAPA_antibiotics,_PABA,_3-hydroxyanthranilate_and_more.	0.5%
Glycolysis_and_Gluconeogenesis,_including_Archaeal_enzymes	0.5%
ABC_transporter_branched-chain_amino_acid_(TC_3.A.1.4.1)	0.5%
Ribonucleotide_reduction	0.5%

	pyrimidine_conversions	0.5%
	V-Type_ATP_synthase	0.5%
	CBSS-350688.3.peg.1509	0.5%
	Pyruvate_metabolism_II:_acetyl-CoA,_acetogenesis_from_pyruvate	0.5%
	Fatty_Acid_Biosynthesis_FASII	0.4%
	Calvin-Benson_cycle	0.4%
	HtrA_and_Sec_secretion	0.4%
	Potassium_homeostasis	0.4%
Function	Carbamoyl-phosphate synthase large chain (EC 6.3.5.5)	0.5%
	Cell division protein FtsH (EC 3.4.24.-)	0.5%
	Excinuclease ABC subunit A	0.5%
	Copper-translocating P-type ATPase (EC 3.6.3.4)	0.5%
	Chaperone protein DnaK	0.4%
	Type I restriction-modification system, restriction subunit R (EC 3.1.21.3)	0.4%
	Glycosyltransferase	0.4%
	Acetyl-coenzyme A synthetase (EC 6.2.1.1)	0.4%
	Isoleucyl-tRNA synthetase (EC 6.1.1.5)	0.4%
	Aspartate aminotransferase (EC 2.6.1.1)	0.4%
	Translation elongation factor G	0.3%
	Leucyl-tRNA synthetase (EC 6.1.1.4)	0.3%
	DNA gyrase subunit A (EC 5.99.1.3)	0.3%
	Excinuclease ABC subunit B	0.3%
	Valyl-tRNA synthetase (EC 6.1.1.9)	0.3%
	UDP-glucose 4-epimerase (EC 5.1.3.2)	0.3%
	DNA polymerase III alpha subunit (EC 2.7.7.7)	0.3%
	DNA-directed RNA polymerase beta' subunit (EC 2.7.7.6)	0.3%
	Type I restriction-modification system, DNA-methyltransferase subunit M (EC 2.1.1.72)	0.3%
	decarboxylase	0.3%
	DNA-directed RNA polymerase beta subunit (EC 2.7.7.6)	0.3%
	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	0.3%
	ClpB protein	0.3%
	CoB--CoM heterodisulfide reductase subunit A (EC 1.8.98.1)	0.3%
	Heat shock protein 60 family chaperone GroEL	0.3%
	Cysteine desulfurase (EC 2.8.1.7)	0.3%
	GTP-binding protein	0.3%
	Threonyl-tRNA synthetase (EC 6.1.1.3)	0.3%
	Pyruvate-flavodoxin oxidoreductase (EC 1.2.7.-)	0.3%
	Alanyl-tRNA synthetase (EC 6.1.1.7)	0.3%

ATP-dependent protease La (EC 3.4.21.53) Type I	0.3%
Glutamate synthase [NADPH] large chain (EC 1.4.1.13)	0.2%
Glucosamine--fructose-6-phosphate aminotransferase [isomerizing] (EC 2.6.1.16)	0.2%
Enolase (EC 4.2.1.11)	0.2%
CTP synthase (EC 6.3.4.2)	0.2%
Protein export cytoplasm protein SecA ATPase RNA helicase (TC 3.A.5.1.1)	0.2%
DNA gyrase subunit B (EC 5.99.1.3)	0.2%
Glutamine synthetase type I (EC 6.3.1.2)	0.2%
DNA topoisomerase I (EC 5.99.1.2)	0.2%
Adenosylhomocysteinase (EC 3.3.1.1)	0.2%
DNA polymerase I (EC 2.7.7.7)	0.2%
Glycogen phosphorylase (EC 2.4.1.1)	0.2%
GMP synthase [glutamine-hydrolyzing] (EC 6.3.5.2)	0.2%
Phenylacetate-coenzyme A ligase (EC 6.2.1.30)	0.2%
Cell division protein FtsZ (EC 3.4.24.-)	0.2%
Ribonucleotide reductase of class III (anaerobic), large subunit (EC 1.17.4.2)	0.2%
Methionyl-tRNA synthetase (EC 6.1.1.10)	0.2%
Ferrous iron transport protein B	0.2%
Acetolactate synthase large subunit (EC 2.2.1.6)	0.2%
Translation initiation factor 2	0.2%

Table S6. Categories of functional genes at different levels based on cDNA sequences from Subsystems database (only top 30 at Level 2, top 50 at Level 3 and Function were listed, Level 1 see Figure 5)

Level 2	NULL	17.4%
	Protein biosynthesis	9.3%
	Central carbohydrate metabolism	6.4%
	Plant-Prokaryote DOE project	4.9%
	Flagellar motility in Prokaryota	2.9%
	ABC transporters	2.9%
	Electron donating reactions	2.8%
	Monosaccharides	2.7%
	Alanine, serine, and glycine	2.4%
	Purines	2.1%
	Lysine, threonine, methionine, and cysteine	2.1%
	Selenoproteins	2.1%
	Fermentation	2.1%
	Oxidative stress	1.9%
	Arginine; urea cycle, polyamines	1.7%
	Transcription	1.7%
	RNA processing and modification	1.7%
	Protein degradation	1.5%
	Pathogenicity islands	1.5%
	ATP synthases	1.3%
	Di- and oligosaccharides	1.3%
	Glutamine, glutamate, aspartate, asparagine; ammonia assimilation	1.2%
	DNA repair	1.2%
	Pyridoxine	1.1%
	One-carbon Metabolism	1.1%
	CO2 fixation	1.1%
	Sugar alcohols	1.0%
	Heat shock	1.0%
	Branched-chain amino acids	0.9%
	Aromatic amino acids and derivatives	0.9%
Level 3	Ribosome_LSU_bacterial	2.5%
	ABC_transporter_oligopeptide_(TC_3.A.1.5.1)	2.1%
	Sugar_utilization_in_Thermotogales	2.0%
	Hydrogenases	1.9%

Glycine_reductase,_sarcosine_reductase_and_betaine_reductase	1.7%
Pyruvate:ferredoxin_oxidoreductase	1.6%
Staphylococcal_pathogenicity_islands_SaPI	1.5%
Ribosome_SSU_bacterial	1.4%
Flagellum_in_Campylobacter	1.3%
RNA_polymerase_bacterial	1.3%
Glycolysis_and_Gluconeogenesis	1.2%
Glycolysis_and_Gluconeogenesis,_including_Archaeal_enzymes	1.2%
Translation_elongation_factors_bacterial	1.2%
Pyridoxin_(Vitamin_B6)_Biosynthesis	1.1%
Translation_elongation_factor_G_family	1.1%
Flagellum	1.1%
Purine_conversions	1.0%
Heat_shock_dnaK_gene_cluster_extended	1.0%
Calvin-Benson_cycle	0.9%
Glycine_and_Serine_Utilization	0.9%
Serine-glyoxylate_cycle	0.9%
Acetone_Butanol_Ethanol_Synthesis	0.9%
Pyruvate_metabolism_I:_anaplerotic_reactions,_PEP	0.8%
PROSC	0.8%
Rubrerythrin	0.7%
HtrA_and_Sec_secretion	0.7%
Glycine_cleavage_system	0.7%
Glutamine,_Glutamate,_Aspartate_and_Asparagine_Biosynthesis	0.7%
V-Type_ATP_synthase	0.7%
Proteolysis_in_bacteria,_ATP-dependent	0.6%
DNA_structural_proteins,_bacterial	0.6%
Polyhydroxybutyrate_metabolism	0.6%
De_Novo_Purine_Biosynthesis	0.6%
F0F1-type_ATP_synthase	0.6%
Glycine_Biosynthesis	0.6%
Bacterial_Cell_Division	0.6%
Anaerobic_Oxidative_Degradation_of_L-Ornithine	0.6%
Flagellar_motility	0.6%
CBSS-312309.3.peg.1965	0.6%
Maltose_and_Maltodextrin_Utilization	0.5%
DNA-replication	0.5%
Oxidative_stress	0.5%
Phosphate_metabolism	0.5%

	Acetyl-CoA_fermentation_to_Butyrate	0.5%
	Purine_Utilization	0.5%
	Cell_division-ribosomal_stress_proteins_cluster	0.5%
	COG1399	0.5%
	Arginine_and_Ornithine_Degradation	0.5%
	Fatty_Acid_Biosynthesis_FASII	0.5%
	Arginine_Biosynthesis_extended	0.4%
Function	Heat shock protein 60 family chaperone GroEL	1.4%
	Pyruvate-flavodoxin oxidoreductase (EC 1.2.7.-)	1.3%
	Flagellin protein FlaA	1.3%
	Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein OppA (TC 3.A.1.5.1)	1.2%
	Translation elongation factor Tu	1.2%
	NAD-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)	0.9%
	Translation elongation factor G	0.8%
	Glycerol kinase (EC 2.7.1.30)	0.7%
	DNA-directed RNA polymerase beta' subunit (EC 2.7.7.6)	0.7%
	Chaperone protein DnaK	0.7%
	Pyruvate,phosphate dikinase (EC 2.7.9.1)	0.6%
	NAD-reducing hydrogenase subunit HoxF (EC 1.12.1.2)	0.6%
	Periplasmic [Fe] hydrogenase large subunit (EC 1.12.7.2)	0.6%
	Phosphoglycerate kinase (EC 2.7.2.3)	0.6%
	DNA-directed RNA polymerase beta subunit (EC 2.7.7.6)	0.6%
	NAD-specific glutamate dehydrogenase (EC 1.4.1.2)	0.5%
	Enolase (EC 4.2.1.11)	0.5%
	Acetyl-CoA acetyltransferase (EC 2.3.1.9)	0.5%
	Flagellin	0.4%
	NADP-specific glutamate dehydrogenase (EC 1.4.1.4)	0.4%
	Rubrerythrin	0.4%
	Fructose-bisphosphate aldolase class II (EC 4.1.2.13)	0.4%
	Ribose ABC transport system, periplasmic ribose-binding protein RbsB (TC 3.A.1.2.1)	0.4%
	Cell division protein FtsH (EC 3.4.24.-)	0.4%
	2-amino-3-ketobutyrate coenzyme A ligase (EC 2.3.1.29)	0.4%
	Aspartate aminotransferase (EC 2.6.1.1)	0.4%
	Glycine reductase component B gamma subunit (EC 1.21.4.2)	0.4%
	DNA-binding protein HBsu	0.4%
	Hydroxylamine reductase (EC 1.7.-.-)	0.4%
	Glycine dehydrogenase [decarboxylating] (glycine cleavage system P2 protein) (EC 1.4.4.2)	0.4%

Glycine reductase component B alpha subunit (EC 1.21.4.2)	0.4%
ClpB protein	0.4%
Glycerol-3-phosphate ABC transporter, periplasmic glycerol-3-phosphate-binding protein (TC 3.A.1.1.3)	0.3%
Glycine reductase component B beta subunit (EC 1.21.4.2)	0.3%
TonB-dependent receptor	0.3%
Oligopeptide transport ATP-binding protein OppD (TC 3.A.1.5.1)	0.3%
Thioredoxin	0.3%
selenocysteine-containing	0.3%
Polyribonucleotide nucleotidyltransferase (EC 2.7.7.8)	0.3%
Phosphoenolpyruvate carboxykinase [GTP] (EC 4.1.1.32)	0.3%
FIG002344: Hydrolase (HAD superfamily)	0.3%
Heat shock protein 60 family co-chaperone GroES	0.3%
V-type ATP synthase subunit B (EC 3.6.3.14)	0.3%
Serine hydroxymethyltransferase (EC 2.1.2.1)	0.3%
Butyryl-CoA dehydrogenase (EC 1.3.99.2)	0.3%
Xanthine/uracil/thiamine/ascorbate permease family protein	0.3%
LSU ribosomal protein L1p (L10Ae)	0.3%
Short chain fatty acids transporter	0.3%
RecA protein	0.3%
Translation elongation factor G-related protein	0.3%

Table S7. Gene functions annotated by COG based on DNA sequences (only top 50 are listed).

Function	Percent
Permeases of the major facilitator superfamily	0.5%
ABC-type multidrug transport system, ATPase component	0.5%
ABC-type multidrug transport system, ATPase and permease components	0.5%
Glycosyltransferases involved in cell wall biogenesis	0.5%
Acyl-coenzyme A synthetases/AMP-(fatty) acid ligases	0.5%
Aspartate/tyrosine/aromatic aminotransferase	0.5%
Isopropylmalate/homocitrate/citramalate synthases	0.5%
ATPases with chaperone activity, ATP-binding subunit	0.4%
Carbamoylphosphate synthase large subunit (split gene in MJ)	0.4%
DNA-directed RNA polymerase, beta' subunit/160 kD subunit	0.4%
Nucleoside-diphosphate-sugar epimerases	0.4%
Translation elongation factors (GTPases)	0.4%
Excinuclease ATPase subunit	0.4%
Phosphoenolpyruvate synthase/pyruvate phosphate dikinase	0.4%
Pyruvate:ferredoxin oxidoreductase and related 2-oxoacid:ferredoxin oxidoreductases, alpha subunit	0.4%
Cation transport ATPase	0.4%
DNA-directed RNA polymerase, beta subunit/140 kD subunit	0.4%
Chaperonin GroEL (HSP60 family)	0.4%
Predicted dehydrogenases and related proteins	0.4%
Glycosyltransferase	0.4%
Dehydrogenases with different specificities (related to short-chain alcohol dehydrogenases)	0.4%
Molecular chaperone	0.4%
SAM-dependent methyltransferases	0.4%
Fe-S oxidoreductase	0.4%
ATPases of the AAA+ class	0.3%
FOG: CBS domain	0.3%
Pyruvate:ferredoxin oxidoreductase and related 2-oxoacid:ferredoxin oxidoreductases, gamma subunit	0.3%
Heterodisulfide reductase, subunit A and related polyferredoxins	0.3%
NAD-dependent aldehyde dehydrogenases	0.3%
Isoleucyl-tRNA synthetase	0.3%
FOG: CheY-like receiver	0.3%
Type IIA topoisomerase (DNA gyrase/topo II, topoisomerase IV), A	0.3%

subunit	
Cation/multidrug efflux pump	0.3%
Leucyl-tRNA synthetase	0.3%
Predicted pyridoxal phosphate-dependent enzyme apparently involved in regulation of cell wall biogenesis	0.3%
Response regulators consisting of a CheY-like receiver domain and a winged-helix DNA-binding domain	0.3%
Response regulator containing CheY-like receiver, AAA-type ATPase, and DNA-binding domains	0.3%
Acetyl-CoA acetyltransferase	0.3%
Type I restriction-modification system methyltransferase subunit	0.3%
Valyl-tRNA synthetase	0.3%
Helicase subunit of the DNA excision repair complex	0.3%
Acyl-CoA synthetases (AMP-forming)/AMP-acid ligases II	0.3%
Superfamily II DNA/RNA helicases, SNF2 family	0.3%
Anaerobic dehydrogenases, typically selenocysteine-containing	0.3%
Cobalamin biosynthesis protein CobN and related Mg-chelatases	0.3%
Thiamine pyrophosphate-requiring enzymes [acetolactate synthase, pyruvate dehydrogenase (cytochrome), glyoxylate carboligase, phosphonopyruvate decarboxylase]	0.3%
Glutamyl- and glutaminyl-tRNA synthetases	0.2%
ATP-dependent Lon protease, bacterial type	0.2%
Pyruvate:ferredoxin oxidoreductase and related 2-oxoacid:ferredoxin oxidoreductases, beta subunit	0.2%
Threonyl-tRNA synthetase	0.2%

Table S8. Gene functions annotated by COG based on cDNA sequences (only top 50 are listed).

Function	Percent
Flagellin and related hook-associated proteins	1.8%
Chaperonin GroEL (HSP60 family)	1.3%
Pyruvate:ferredoxin oxidoreductase and related 2-oxoacid:ferredoxin oxidoreductases, alpha subunit	1.3%
ABC-type sugar transport system, periplasmic component	1.1%
GTPases - translation elongation factors	1.1%
Glutamate dehydrogenase/leucine dehydrogenase	1.0%
Translation elongation factors (GTPases)	1.0%
Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase	0.9%
ATPases with chaperone activity, ATP-binding subunit	0.8%
Phosphoenolpyruvate synthase/pyruvate phosphate dikinase	0.8%
Pyruvate:ferredoxin oxidoreductase and related 2-oxoacid:ferredoxin oxidoreductases, gamma subunit	0.8%
ABC-type oligopeptide transport system, periplasmic component	0.7%
Uncharacterized ABC-type transport system, periplasmic component/surface lipoprotein	0.7%
DNA-directed RNA polymerase, beta' subunit/160 kD subunit	0.7%
Acetyl-CoA acetyltransferase	0.7%
Molecular chaperone	0.7%
Glycerol kinase	0.7%
Bacterial nucleoid DNA-binding protein	0.7%
Predicted dehydrogenases and related proteins	0.6%
NADH:ubiquinone oxidoreductase, NADH-binding (51 kD) subunit	0.6%
3-phosphoglycerate kinase	0.5%
Enolase	0.5%
DNA-directed RNA polymerase, beta subunit/140 kD subunit	0.5%
Response regulators consisting of a CheY-like receiver domain and a winged-helix DNA-binding domain	0.5%
ABC-type dipeptide transport system, periplasmic component	0.5%
ABC-type sugar transport systems, ATPase components	0.5%
NAD-dependent aldehyde dehydrogenases	0.5%
Predicted HD superfamily hydrolase	0.4%
Pyruvate:ferredoxin oxidoreductase and related 2-oxoacid:ferredoxin oxidoreductases, beta subunit	0.4%
Aspartate/tyrosine/aromatic aminotransferase	0.4%

Acyl-CoA dehydrogenases	0.4%
Fructose/tagatose bisphosphate aldolase	0.4%
Rubrerythrin	0.4%
7-keto-8-aminopelargonate synthetase and related enzymes	0.4%
NADH dehydrogenase/NADH:ubiquinone oxidoreductase 75 kD subunit (chain G)	0.4%
Molecular chaperone (small heat shock protein)	0.4%
Outer membrane protein and related peptidoglycan-associated (lipo)proteins	0.4%
ATP-dependent Zn proteases	0.4%
Dehydrogenases with different specificities (related to short-chain alcohol dehydrogenases)	0.4%
Predicted cobalamin binding protein	0.3%
Superoxide dismutase	0.3%
Glycine cleavage system protein P (pyridoxal-binding), C-terminal domain	0.3%
DNA-directed RNA polymerase, sigma subunit (sigma70/sigma32)	0.3%
ABC-type oligopeptide transport system, ATPase component	0.3%
Predicted pyridoxal phosphate-dependent enzyme apparently involved in regulation of cell wall biogenesis	0.3%
6Fe-6S prismane cluster-containing protein	0.3%
Phosphoenolpyruvate carboxykinase (GTP)	0.3%
Acyl CoA:acetate/3-ketoacid CoA transferase, alpha subunit	0.3%
Predicted Fe-S oxidoreductases	0.3%
Polyribonucleotide nucleotidyltransferase (polynucleotide phosphorylase)	0.3%

Table S9. COG categories of functional genes of DNA, cDNA datasets and selected microbes.

Microbiomes	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Metagenome	0.0	0.1	11.0	1.0	10.4	4.3	7.2	5.8	3.0	10.1	3.2	7.7	5.1	0.2	5.2	5.3	1.3	12.2	0.5	2.4	1.7	2.4	0.0	0.0	0.0	0.0
Metatranscriptome	0.0	0.0	14.0	0.8	11.3	3.3	12.0	2.6	3.9	12.6	4.1	4.8	3.9	2.2	6.7	2.5	0.9	9.7	0.4	1.8	1.6	1.0	0.0	0.0	0.0	0.0
Clostridium kluyveri	0.0	0.0	5.8	2.1	9.9	2.7	4.0	4.8	3.3	8.2	7.7	5.0	5.8	2.9	4.3	6.4	2.2	6.9	5.6	5.5	0.9	3.7	0.0	2.4	0.0	0.0
Ruminococcaceae bacterium	0.0	0.0	6.8	1.7	9.5	3.2	7.0	4.5	2.8	10.6	8.1	7.0	5.4	0.3	3.2	4.9	0.8	7.1	4.7	5.5	1.7	3.6	0.0	1.4	0.0	0.0
Petrotoga mobilis	0.0	0.0	6.5	1.3	12.9	4.6	10.0	4.0	3.2	13.7	5.1	5.3	6.2	3.3	4.8	5.0	0.7	7.6	3.5	2.4	0.7	4.5	0.0	1.0	0.0	0.0
Clostridium botulinum	0.0	0.0	6.0	2.2	8.6	4.1	6.7	5.0	3.4	10.3	7.1	5.5	5.8	3.8	4.3	5.1	0.9	6.8	5.2	4.3	0.7	2.7	0.0	1.5	0.0	0.0
Clostridium thermocellum	0.0	0.0	5.7	2.6	8.1	3.1	7.4	4.6	2.2	10.2	5.4	5.4	7.1	4.3	4.8	4.1	0.9	7.2	4.9	5.1	0.8	2.9	0.0	3.2	0.0	0.1
Clostridium difficile	0.0	0.0	6.6	1.4	9.4	3.5	9.7	4.1	2.6	8.8	9.5	5.3	5.9	2.0	3.6	4.7	1.2	5.9	5.6	5.1	0.8	3.1	0.0	1.0	0.0	0.1
Clostridium carboxidivorans	0.0	0.0	8.0	1.6	10.5	2.6	5.8	4.4	2.9	7.3	9.4	3.9	5.7	3.4	3.9	4.7	1.9	7.0	5.9	6.5	0.4	3.7	0.0	0.5	0.0	0.0
Clostridium cellulolyticum	0.0	0.0	5.1	2.4	8.1	3.0	9.1	5.2	3.8	9.5	6.0	6.2	5.5	3.2	4.0	3.8	2.0	7.1	5.0	5.1	1.0	2.9	0.0	2.1	0.0	0.0
Clostridium acetobutylicum	0.0	0.0	5.6	1.7	8.2	2.9	11.4	4.9	3.2	7.2	9.4	4.1	6.2	3.1	3.9	4.5	1.1	7.0	5.9	5.4	0.5	2.9	0.0	0.8	0.0	0.0
Alkaliphilus metalliredigens	0.0	0.0	8.1	1.6	12.0	3.0	5.9	4.2	2.3	7.7	6.7	4.5	5.2	2.7	3.8	5.3	1.6	6.7	6.2	5.6	0.7	4.0	0.0	2.0	0.0	0.0
Alkaliphilus oremlandii	0.0	0.0	7.6	2.2	8.9	3.6	2.9	5.2	3.3	10.5	7.4	5.2	5.1	3.0	4.1	6.6	1.7	7.1	6.1	5.1	0.8	3.1	0.0	0.5	0.0	0.1
Bacteroides sp	0.0	0.0	5.4	1.0	9.4	6.5	12.4	5.0	3.5	8.8	4.2	5.2	10.2	0.3	4.2	4.2	0.6	6.4	4.0	4.0	0.9	2.9	0.0	0.9	0.0	0.0
Bacteroides fragilis	0.0	0.0	6.5	1.2	8.0	3.5	9.5	5.3	2.9	9.2	4.0	5.8	12.0	0.3	4.4	6.0	0.8	6.9	3.9	4.2	1.1	3.9	0.0	0.5	0.0	0.0
Methanosaeta thermophi	0.2	0.5	8.3	0.8	7.7	4.5	3.4	8.9	1.8	14.1	4.7	4.8	4.1	0.1	4.7	6.5	0.7	10.4	7.9	2.6	0.8	2.3	0.5	0.5	0.0	0.0
Methanobrevibacter smithii	0.1	0.2	9.6	0.6	7.7	4.3	2.8	9.7	2.5	13.7	4.9	5.4	4.6	0.3	4.5	4.3	0.7	9.9	9.0	0.9	0.9	2.6	0.1	0.7	0.0	0.0
Methanoculleus marisnigri	0.1	0.2	8.8	0.7	6.9	3.6	3.8	8.1	1.9	10.6	5.2	4.9	4.6	1.3	5.7	5.7	0.8	11.0	8.4	4.0	0.7	2.6	0.1	0.2	0.0	0.0
Methanoregula boonei	0.0	0.1	9.6	0.8	7.0	3.3	2.9	7.5	1.9	10.5	5.5	4.7	2.6	1.1	4.7	6.3	0.9	9.5	8.7	8.2	0.9	2.5	0.4	0.2	0.0	0.0
Methanosphaera stadtmanae	0.0	0.5	9.2	0.7	8.3	4.3	3.3	9.8	2.2	14.9	4.5	5.4	5.6	0.3	4.3	4.4	0.3	8.8	8.0	1.3	1.0	2.8	0.2	0.1	0.0	0.0
Methanothermobacter marburgensis	0.1	0.3	10.5	0.4	6.3	4.5	3.1	9.3	1.9	12.7	4.5	4.5	2.9	0.3	5.0	5.5	0.9	11.1	10.0	3.5	1.0	1.3	0.1	0.1	0.0	0.0
Methanothermobacter thermautotrophicus	0.1	0.3	10.7	0.4	5.7	4.2	3.0	9.0	2.4	11.8	4.3	4.3	3.6	0.3	5.5	5.7	0.6	11.5	10.1	3.5	0.9	2.1	0.1	0.0	0.0	0.0

Note: A: RNA processing and modification; B: Chromatin structure and dynamics; C: Energy production and conversion; D: Cell cycle control, cell division, chromosome partitioning; E: Amino acid transport and metabolism; F: Nucleotide transport and metabolism; G: Carbohydrate transport and metabolism; H: Coenzyme transport and metabolism; I: Lipid transport and metabolism; J: Lipid transport and metabolism; K: Transcription; L: Replication, recombination and repair; M: Cell wall/membrane/envelope biogenesis; N: Cell motility; 0: Posttranslational modification, protein turnover, chaperones; P: Inorganic ion transport and metabolism; Q: Secondary metabolites biosynthesis, transport and catabolism; R: General function prediction only; S: Function unknown; T: Signal transduction mechanisms; U: Intracellular trafficking, secretion, and vesicular transport; V: Defense mechanisms; W: Extracellular structures; X: Mobilome: prophages, transposons; Y: Nuclear structure; Z: Cytoskelet