Table S2. COG annotation of the 97 proteins used for phylogenomic analysis.

COG id	Biological function
COG0541	Signal recognition particle GTPase
COG2511	Archaeal Glu-tRNAGln amidotransferase subunit E (contains GAD domain)
COG0252	L-asparaginase/archaeal Glu-tRNAGln amidotransferase subunit D
COG5257	Translation initiation factor 2, gamma subunit (eIF-2gamma; GTPase)
COG0152	Phosphoribosylaminoimidazolesuccinocarboxamide (SAICAR) synthase
COG1797	Cobyrinic acid a,c-diamide synthase
COG2138	Uncharacterized conserved protein
COG2082	Precorrin isomerase
COG0096	Ribosomal protein S8
COG0113	Delta-aminolevulinic acid dehydratase
COG1881	Phospholipid-binding protein
COG2109	ATP:corrinoid adenosyltransferase
COG0468	RecA/RadA recombinase
COG0126	3-phosphoglycerate kinase
COG0520	Selenocysteine lyase
COG1093	Translation initiation factor 2, alpha subunit (eIF-2alpha)
COG0615	Cytidylyltransferase
COG0097	Ribosomal protein L6P/L9E
COG0128	5-enolpyruvylshikimate-3-phosphate synthase
COG2260	Predicted Zn-ribbon RNA-binding protein
COG0093	Ribosomal protein L14
COG0396	ABC-type transport system involved in Fe-S cluster assembly, ATPase component
COG0090	Ribosomal protein L2
COG0092	Ribosomal protein S3
COG1471	Ribosomal protein S4E
COG0034	Glutamine phosphoribosylpyrophosphate amidotransferase
COG1339	Transcriptional regulator of a riboflavin/FAD biosynthetic operon
COG2090	Uncharacterized protein conserved in archaea
COG1675	Transcription initiation factor IIE, alpha subunit
COG0048	Ribosomal protein S12
COG2125	Ribosomal protein S6E (S10)
COG0225	Peptide methionine sulfoxide reductase
COG0504	CTP synthase (UTP-ammonia lyase)
COG0169	Shikimate 5-dehydrogenase
COG1303	Uncharacterized protein conserved in archaea
COG2139	Ribosomal protein L21E
COG1324	Uncharacterized protein involved in tolerance to divalent cations
COG2262	GTPases
COG0088	Ribosomal protein L4
COG0100	Ribosomal protein S11
COG0010	Arginase/agmatinase/formimionoglutamate hydrolase, arginase family
COG1958	Small nuclear ribonucleoprotein (snRNP) homolog
COG1646	Predicted phosphate-binding enzymes, TIM-barrel fold
COG0189	Glutathione synthase/Ribosomal protein S6 modification enzyme (glutaminyl transferase)
COG4830	Ribosomal protein S26
COG1547	Uncharacterized conserved protein
COG1258	Predicted pseudouridylate synthase
COG0265	Trypsin-like serine proteases, typically periplasmic, contain C-terminal PDZ domain

COG1412	Uncharacterized proteins of PilT N-term./Vapc superfamily
COG0199	Ribosomal protein S14
COG0287	Prephenate dehydrogenase
COG1798	Diphthamide biosynthesis methyltransferase
COG0094	Ribosomal protein L5
COG1903	Cobalamin biosynthesis protein CbiD
COG2890	Methylase of polypeptide chain release factors
COG2073	Cobalamin biosynthesis protein CbiG
COG1867	N2,N2-dimethylguanosine tRNA methyltransferase
COG2875	Precorrin-4 methylase
COG0667	Predicted oxidoreductases (related to aryl-alcohol dehydrogenases)
COG1491	Predicted RNA-binding protein
COG2429	Uncharacterized conserved protein
COG1985	Pyrimidine reductase, riboflavin biosynthesis
COG0358	DNA primase (bacterial type)
COG0057	Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase
COG1024	Enoyl-CoA hydratase/carnithine racemase
COG1537	Predicted RNA-binding proteins
COG0671	Membrane-associated phospholipid phosphatase
COG4221	Short-chain alcohol dehydrogenase of unknown specificity
COG0087	Ribosomal protein L3
COG0186	Ribosomal protein S17
COG1460	Uncharacterized protein conserved in archaea
COG0091	Ribosomal protein L22
COG0644	Dehydrogenases (flavoproteins)
COG1587	Uroporphyrinogen-III synthase
COG0049	Ribosomal protein S7
COG0030	Dimethyladenosine transferase (rRNA methylation)
COG2241	Precorrin-6B methylase 1
COG0185	Ribosomal protein S19
COG0863	DNA modification methylase
COG1180	Pyruvate-formate lyase-activating enzyme
COG0195	Transcription elongation factor
COG3253	Uncharacterized conserved protein
COG1254	Acylphosphatases
COG1378	Predicted transcriptional regulators
COG1439	Predicted nucleic acid-binding protein, consists of a PIN domain and a Zn-ribbon module
COG1522	Transcriptional regulators
COG1964	Predicted Fe-S oxidoreductases
COG1599	Single-stranded DNA-binding replication protein A (RPA), large (70 kD) subunit and related ssDNA-binding proteins
COG1940	Transcriptional regulator/sugar kinase
COG0054	Riboflavin synthase beta-chain
COG0255	Ribosomal protein L29
COG2242	Precorrin-6B methylase 2
COG1409	Predicted phosphohydrolases
COG1382	Prefoldin, chaperonin cofactor
COG1703	Putative periplasmic protein kinase ArgK and related GTPases of G3E family
COG3185	4-hydroxyphenylpyruvate dioxygenase and related hemolysins
COG0805	Sec-independent protein secretion pathway component TatC