

Table 6. The 60 most frequently occurring COG neighborhoods unique to metagenomic datasets

Occurrences	COG A	COG B	COG A Description	COG B Description
171	COG1077	COG0119	Actin-like ATPase involved in cell morphogenesis	Isopropylmalate/homocitrate/citramalate synthases
161	COG0015	COG0024	Adenylosuccinate lyase	Methionine aminopeptidase
153	COG0190	COG0762	5,10-methylene-tetrahydrofolate dehydrogenase/Methenyl tetrahydrofolate cyclohydrolase	Predicted integral membrane protein
149	COG0072	COG0481	Phenylalanyl-tRNA synthetase beta subunit	Membrane GTPase LepA
144	COG0527	COG0821	Aspartokinases	Enzyme involved in the deoxyxylulose pathway of isoprenoid biosynthesis
143	COG2086	COG0404	Electron transfer flavoprotein, beta subunit	Glycine cleavage system T protein (aminomethyltransferase)
141	COG0141	COG0361	Histidinol dehydrogenase	Translation initiation factor 1 (IF-1)
139	COG0008	COG1974	Glutamyl- and glutaminyl-tRNA synthetases	SOS-response transcriptional repressors (RecA-mediated autopeptidases)
134	COG0289	COG0177	Dihydrodipicolinate reductase	Predicted EndoIII-related endonuclease
133	COG0192	COG0815	S-adenosylmethionine synthetase	Apolipoprotein N-acyltransferase
129	COG0119	COG0215	Isopropylmalate/homocitrate/citramalate	CysteinyI-tRNA synthetase
128	COG0524	COG0177	Sugar kinases, ribokinase family	Predicted EndoIII-related endonuclease
126	COG0483	COG0254	Archaeal fructose-1,6-bisphosphatase and related enzymes of inositol monophosphatase	Ribosomal protein L31
124	COG0188	COG0629	Type IIA topoisomerase (DNA gyrase/topo II, topoisomerase IV), A subunit	Single-stranded DNA-binding protein
123	COG0254	COG3820	Ribosomal protein L31	Uncharacterized protein conserved in bacteria
123	COG0192	COG0195	S-adenosylmethionine synthetase	Transcription elongation factor
122	COG0477	COG0168	Permeases of the major facilitator superfamily	Trk-type K+ transport systems, membrane components
121	COG0349	COG0794	Ribonuclease D	Predicted sugar phosphate isomerase involved in capsule formation
121	COG0206	COG2001	Cell division GTPase	Uncharacterized protein conserved in bacteria
120	COG1304	COG1304	L-lactate dehydrogenase (FMN-dependent) and related alpha-hydroxy acid dUTPase	L-lactate dehydrogenase (FMN-dependent) and related alpha-hydroxy acid
120	COG0756	COG0476	dUTPase	Dinucleotide-utilizing enzymes involved in molybdopterin and thiamine biosynthesis
119	COG0167	COG1384	Dihydroorotate dehydrogenase	Lysyl-tRNA synthetase (class I)
117	COG0668	COG0762	Small-conductance mechanosensitive channel	Predicted integral membrane protein
117	COG0499	COG0802	S-adenosylhomocysteine hydrolase	Predicted ATPase or kinase
116	COG0604	COG1741	NADPH:quinone reductase and related Zn-dependent oxidoreductases	Pirip-related protein
115	COG0462	COG1496	Phosphoribosylpyrophosphate synthetase	Uncharacterized conserved protein
115	COG0408	COG0782	Coproporphyrinogen III oxidase	Transcription elongation factor
114	COG0400	COG1403	Predicted esterase	Restriction endonuclease
114	COG0104	COG0351	Adenylosuccinate synthase	Hydroxymethylpyrimidine/phosphomethylpyrimidine kinase
113	COG0180	COG0694	Tryptophanyl-tRNA synthetase	Thioredoxin-like proteins and domains
112	COG0225	COG2897	Peptide methionine sulfoxide reductase	Rhodanese-related sulfurtransferase
112	COG0408	COG0219	Coproporphyrinogen III oxidase	Predicted rRNA methylase (SpoU class)
112	COG2009	COG0136	Succinate dehydrogenase/fumarate reductase, cytochrome b subunit	Aspartate-semialdehyde dehydrogenase
109	COG0119	COG0565	Isopropylmalate/homocitrate/citramalate synthases	rRNA methylase
108	COG0793	COG0799	Periplasmic protease	Uncharacterized homolog of plant lojap
108	COG0717	COG0766	Deoxycytidine deaminase	UDP-N-acetylglucosamine enolpyruvyl
107	COG1304	COG0436	L-lactate dehydrogenase (FMN-dependent) and related alpha-hydroxy acid	Aspartate/tyrosine/aromatic aminotransferase
105	COG0343	COG0652	Queuine/archaeosine tRNA-ribosyltransferase	Peptidyl-prolyl cis-trans isomerase (rotamase) : cyclophilin family
104	COG0289	COG0484	Dihydrodipicolinate reductase	DnaJ-class molecular chaperone with C-terminal Zn finger domain
102	COG0492	COG0625	Thioredoxin reductase	Glutathione S-transferase
101	COG0686	COG0161	Alanine dehydrogenase	Adenosylmethionine-8-amino-7-oxononanoate aminotransferase
101	COG0006	COG3473	Xaa-Pro aminopeptidase	Maleate cis-trans isomerase
100	COG621	COG0735	2-methylthioadenine synthetase	Fe2+/Zn2+ uptake regulation proteins
100	COG1109	COG0351	Phosphomannomutase	Hydroxymethylpyrimidine/phosphomethylpyrimidine kinase
98	COG0461	COG0854	Orotate phosphoribosyltransferase	Pyridoxal phosphate biosynthesis protein
97	COG0821	COG0216	Enzyme involved in the deoxyxylulose pathway of isoprenoid biosynthesis	Protein chain release factor A
96	COG0077	COG2812	Prephenate dehydratase	DNA polymerase III, gamma/tau subunits
95	COG0548	COG0706	Acetylglutamate kinase	Preprotein translocase subunit YidC
95	COG0479	COG1012	Succinate dehydrogenase/fumarate reductase, Fe-S protein subunit	NAD-dependent aldehyde dehydrogenases
94	COG1028	COG0861	Dehydrogenases with different specificities (related to short-chain alcohol)	Membrane protein TerC, possibly involved in tellurium resistance
94	COG0284	COG0776	Orotidine-5-phosphate decarboxylase	Bacterial nucleoid DNA-binding protein
94	COG0134	COG1974	Indole-3-glycerol phosphate synthase	SOS-response transcriptional repressors (RecA-mediated autopeptidases)
93	COG0129	COG0525	Dihydroxyacid dehydratase/phosphogluconate dehydratase	Valyl-tRNA synthetase
92	COG0324	COG0265	tRNA delta(2)-isopentenylpyrophosphate transferase	Trypsin-like serine proteases, typically periplasmic, contain C-terminal PDZ domain
91	COG0604	COG0302	NADPH:quinone reductase and related Zn-dependent oxidoreductases	GTP cyclohydrolase I
91	COG0128	COG1137	5-enolpyruvylshikimate-3-phosphate synthase	ABC-type (unclassified) transport system, ATPase component
91	COG0006	COG0210	Xaa-Pro aminopeptidase	Superfamily I DNA and RNA helicases
89	COG1003	COG2907	Glycine cleavage system protein P (pyridoxal-binding), C-terminal domain	Predicted NAD/FAD-binding protein
89	COG2303	COG0861	Choline dehydrogenase and related flavoproteins	Membrane protein TerC, possibly involved in tellurium resistance
88	COG0382	COG1132	4-hydroxybenzoate polyprenyltransferase and related prenyltransferases	ABC-type multidrug transport system, ATPase and permease components