





COG categories are as follows: For cellular processes and signaling, D is cell cycle control, cell division, and chromosome partitioning; M is cell wall/membrane/envelope biogenesis; N is cell motility; O is posttranslational modification, protein turnover, and chaperones; T is signal transduction mechanisms; U is intracellular trafficking, secretion, and vesicular transport; V is defense mechanisms; W is extracellular structures; Y is nuclear structure; and Z is cytoskeleton. For information storage and processing, A is RNA processing and modification; B is chromatin structure and dynamics; J is translation, ribosomal structure, and biogenesis; K is transcription; and L is replication, recombination, and repair. For metabolism, C is energy production and conversion; E is amino acid transport and metabolism; F is nucleotide transport and metabolism; G is carbohydrate transport and metabolism; H is coenzyme transport and metabolism; I is lipid transport and metabolism; P is inorganic ion transport and metabolism; and Q is secondary metabolites biosynthesis, transport, and catabolism. For poorly characterized genes, R is for general function prediction only; and S is for unknown function. Genes belonging to W, Y, Z, A, and B are not found.