The ISME Journal (2017), 1–16

www.nature.com/ismej

ORIGINAL ARTICLE

New insights into marine group III Euryarchaeota, from dark to light

Jose M Haro-Moreno[1,3], Francisco Rodriguez-Valera[1], Purificación López-García[2], David Moreira[2] and Ana-Belen Martin-Cuadrado[1,3]

[1]*Evolutionary Genomics Group, Departamento de Producción Vegetal y Microbiología, Universidad Miguel Hernández, Alicante, Spain and* [2]*Unité d'Ecologie, Systématique et Evolution, UMR CNRS 8079, Université Paris-Sud, Orsay Cedex, France*

Marine Euryarchaeota remain among the least understood major components of marine microbial communities. Marine group II Euryarchaeota (MG-II) are more abundant in surface waters (4–20% of the total prokaryotic community), whereas marine group III Euryarchaeota (MG-III) are generally considered low-abundance members of deep mesopelagic and bathypelagic communities. Using genome assembly from direct metagenome reads and metagenomic fosmid clones, we have identified six novel MG-III genome sequence bins from the photic zone (Epi1–6) and two novel bins from deep-sea samples (Bathy1–2). Genome completeness in those genome bins varies from 44% to 85%. Photic-zone MG-III bins corresponded to novel groups with no similarity, and significantly lower GC content, when compared with previously described deep-MG-III genome bins. As found in many other epipelagic microorganisms, photic-zone MG-III bins contained numerous photolyase and rhodopsin genes, as well as genes for peptide and lipid uptake and degradation, suggesting a photoheterotrophic lifestyle. Phylogenetic analysis of these photolyases and rhodopsins as well as their genomic context suggests that these genes are of bacterial origin, supporting the hypothesis of an MG-III ancestor that lived in the dark ocean. Epipelagic MG-III occur sporadically and in relatively small proportions in marine plankton, representing only up to 0.6% of the total microbial community reads in metagenomes. None of the reconstructed epipelagic MG-III genomes were present in metagenomes from aphotic zone depths or from high latitude regions. Most low-GC bins were highly enriched at the deep chlorophyll maximum zones, with the exception of Epi1, which appeared evenly distributed throughout the photic zone worldwide.

The ISME Journal advance online publication, 13 January 2017; doi:10.1038/ismej.2016.188

Introduction

Marine archaea are important marine microbes in terms of their metabolic activity and abundance (Karner *et al.*, 2001; Li *et al.*, 2015). Ammonia-oxidizing Thaumarchaeota (Brochier-Armanet *et al.*, 2008) are the most abundant archaeal phylum in the oceans and have a key role in the marine nitrogen cycle (Konneke *et al.*, 2005; Qin *et al.*, 2014). Studies have also identified three major groups of marine Euryarchaeota: (i) group II (MG-II) (DeLong, 1992; Fuhrman *et al.*, 1992; Fuhrman and Davis, 1997; Massana *et al.*, 2000), (ii) group III (MG-III) (Fuhrman and Davis, 1997; Lopez-Garcia *et al.*, 2001a), and (iii) group IV (MG-IV) (Lopez-Garcia *et al.*, 2001b). So far, there are no cultured representatives of marine Euryarchaeota and little is known about their physiology and ecological role in the oceans. MG-II are widely distributed within the euphotic zone of temperate waters. MG-II are the dominant archaeal community not only in the surface and in the deep chlorophyll maximum (DCM) (Massana *et al.*, 2000; Karner *et al.*, 2001; Herndl *et al.*, 2005; DeLong *et al.*, 2006; Galand *et al.*, 2010; Belmar *et al.*, 2011; Martin-Cuadrado *et al.*, 2015) but have also been found in deep-sea waters (Lopez-Garcia *et al.*, 2001a,Martin-Cuadrado *et al.*, 2008; Li *et al.*, 2015). The other two marine Euryarchaeota groups, MG-III and MG-IV, are considered to be rare components of deep-sea communities (Lopez-Garcia *et al.*, 2001a,b; Galand *et al.*, 2009).

MG-III were first described by Fuhrman and Davis, 1997 from deep marine plankton samples and have subsequently been found in 16S-rRNA gene surveys from most deep oceanic regions, albeit at very low abundance (Massana *et al.*, 2000; Lopez-Garcia *et al.*, 2001a,b), and by metagenomics throughout the water column in the central Pacific gyre (DeLong *et al.*, 2006).

Correspondence: A-B Martin-Cuadrado, Departamento de Producción Vegetal y Microbiología, Universidad Miguel Hernández, APARTADO 18, CAMPUS SAN JUAN, CP. 03550, San Juan de Alicante, Alicante 03550, Spain.
E-mail: amartin@umh.es
[3]These authors contributed equally to this work.
Received 9 May 2016; revised 25 November 2016; accepted 5 December 2016

However, occasionally, they have been identified at much higher proportions. For instance, 16S-rRNA sequences from MG-III represented one of the largest archaeal groups in the deep Arctic Ocean (>40% of tag sequences) (Galand *et al.*, 2009) and between 30% and 50% of the archaeal sequences from a deep (500–1250 m) Marmara Sea metagenome (Quaiser *et al.*, 2011). They were also relatively abundant (ca.18% of the total archaeal population) in the oxygen minimum zone (50–400 m) in the Eastern tropical South Pacific (Belmar *et al.*, 2011). Only a few studies report the presence of MG-III in the photic zone. They represented 0.4% of all the archaeal sequences obtained in surface Arctic waters (Galand *et al.*, 2009) and up to 10% in samples recovered along 4.5 years in the Mediterranean DCM (Galand *et al.*, 2010).

The initial analysis of three MG-III fosmids from deep-sea metagenomic libraries allowed a first glance at their metabolic potential (Martin-Cuadrado *et al.*, 2008). The presence of some fermentation-related genes led to the hypothesis that they could be facultative anaerobes. In a recent study, up to 3% of the single amplified genomes of archaea recovered from mesopelagic waters from South-Atlantic and North-Pacific gyres belonged to MG-III (Swan *et al.*, 2014). However, only two single amplified genomes classified as MG-III, SCGC-AAA-007-O11 (isolated at 800 m in the South-Atlantic sub-tropical gyre) and SCGC-AAA-288-E19 (from 770 m in the North-Pacific sub-tropical gyre), have been deposited in GenBank. Only the SCGC-AAA-288-E19 partial genome had ribosomal RNA genes that corresponded to MG-III, but contig annotation showed contamination with *Chloroflexi* (32 genome fragments out of the 102). Complete archaeal fosmids (452 adding up to 16 Mb of sequence) from deep Mediterranean samples belonging to MG-II/III have been published (Deschamps *et al.*, 2014) and five MG-III partial genomes (31–65% completeness) were assembled from metagenomes from the Guaymas basin (1993 m, Gulf of California) and the Mid-Cayman Rise (2040–2238 m and 4869–4946 m, Caribbean Sea) (Li *et al.*, 2015). Based on the genes present in these genomes, it was proposed that the microbes they represented are motile heterotrophs with different mechanisms for scavenging organic matter.

Binning the assembled fragments by oligonucleotide frequencies, GC content and differential recruitment in metagenomes is a successful strategy for the discovery of novel microbial lineages (Tyson *et al.*, 2004; Ghai *et al.*, 2012; Iverson *et al.*, 2012; Narasingarao *et al.*, 2012; Martin-Cuadrado *et al.*, 2015; Li *et al.*, 2015; Vavourakis *et al.*, 2016). We applied this approach to recover MG-III sequences using several metagenomic fosmid libraries from the Mediterranean Sea (collections KM3, AD1000 (Martin-Cuadrado *et al.*, 2008) and MedDCM-OCT2007 (Ghai *et al.*, 2010)) and from the assemblies of 16 metagenomes (four collections from the Mediterranean: MedDCM-JUL2012 (Martin-Cuadrado *et al.*, 2015), MedDCM-SEP2014 (this work), Med-Io7–77mDCM and Med-Ae2–600mDeep (Mizuno *et al.*, 2016) and 12 from TARA microbiomes (Sunagawa *et al.*, 2015)). We obtained a total of eight different MG-III genome bins. Six of them belong to novel surface MG-III lineages distantly related to the previously described deep MG-III sequence bins (Li *et al.*, 2015). They are the first near-complete genomes of MG-III living in the photic zone. Some of them appear to be widespread in the ocean; their distribution in different water masses has been analyzed.

Materials and methods

Sampling and sequencing

A fosmid metagenomic library of ca. 13 000 clones was constructed with biomass recovered in October 2007 (50 m deep) at the Mediterranean DCM (38°4′6.64″N 0° 13′55.18″W). Partial results of almost 7000 fosmid sequences have been described previously in Ghai *et al.* (2010) and Martin-Cuadrado *et al.* (2015). Metagenomes were also sequenced from samples recovered at the same location and at a similar depth the following years (MedDCM-JUL2012 (Martin-Cuadrado *et al.*, 2015) and MedDCM-SEP2014) from one sample recovered at the DCM from the Ionian Sea (Med-Io7–77mDCM) and from a sample collected from the deep Aegean Sea (Med-Ae2–600mDeep) (Mizuno *et al.*, 2016). For these metagenomes, sea water was collected and sequentially filtered on-board using a positive pressure system through a 20 μm pore filter followed by a 5 μm pore size polycarbonate filter and, finally, 0.22 μm pore size Sterivex filters (Durapore; Millipore, Billerica, MA, USA). Filters were frozen on dry ice and kept at − 80 °C until processed in the laboratory. Filters were thawed on ice and then treated with 1 mg ml^{-1} lysozyme and 0.2 mg ml^{-1} proteinase K (final concentrations). Nucleic acids were extracted with phenol–chloroform–isoamyl alcohol and chloroform–isoamyl alcohol. Sequencing was carried out using Illumina HiSeq2000 (PE, 100 bp) (Macrogen, Seoul, Korea and BGI, Hong Kong).

'De novo' assembly, gene annotation and binning of the MG-III sequences

A schematic of the assembly pipeline is shown in Supplementary Figure S1. The assembly of the fosmids from the MedDCM-OCT2007, KM3 and AD1000 metagenomic fosmid libraries has been previously described (Ghai *et al.*, 2010; Deschamps *et al.*, 2014; Martin-Cuadrado *et al.*, 2015). Sequences from metagenomes MedDCM-JUL2012, MedDCM-SEP2014, Med-Io7–77mDCM and Med-Ae2–600mDeep were quality trimmed and assembled independently using IBDA-UD (Peng *et al.*, 2012) with the following parameters: –mink 70, –maxk 100, –step 10, –pre_correction. Gene predictions on the assembled sequences were carried out using Prodigal (Hyatt *et al.*, 2010). Ribosomal genes were identified using ssu-align (Nawrocki, 2009) and meta_rna (Huang *et al.*, 2009). Functional annotation was performed by comparing predicted protein

Table 1 General features of the MG-III bins and the composite genomes CG-MGIII

MG-III bin	*No. of contigs*	*%GC ± s.d.*	*No. of Mb*	*% Genome (1) (2) (3)*[a]	*No. of genomes*	*Largest contig (Kb)*	*No. of CDS/nr CDS*[b]	*Intergenic region (bp)*	*Median gene size (bp)*	*CheckM*[c] *% contamination*
Epi1	136	36.6 ± 0.9	2.95	80.0/35.1/83.0	2	120.8	3058/1307	42	752	—
CG-Epi1	25	36.6 ± 0.8	1.18	85.7/34.2/84.9	1	135.9	—/1106	40	792	2
Epi2A	34	36.0 ± 0.9	0.54	54.3/16.2/45.3	1	27.4	519/—	41	765	—
Epi2B	26	36.2 ± 0.9	0.56	57.1/18.0/54.7	1	79.1	525/—	40	768	—
Epi2C	10	36.1 ± 0.9	0.31	5.7/4.5/5.7	1	92.8	282/—	40	807	—
CG-Epi2	43	36.1 ± 1.1	1.22	80.0/30.6/75.5	1	101.9	—/1182	42	777	25
Epi3	35	35.9 ± 1.2	0.71	48.6/19.8/45.3	1	55.8	675/673	37	798	—
CG-Epi3	30	35.9 ± 1.2	0.79	62.9/24.3/62.3	1	71.7	—/741	40	786	0.8
Epi4	31	36.5 ± 0.8	0.66	34.3/15.3/39.6	1	47.2	612/610	37	794	—
CG-Epi4	24	36.4 ± 0.7	0.71	34.3/16.2/43.4	1	80	—/666	37	786	0
Epi5	18	36.1 ± 0.8	0.26	2.9/3.6/7.6	1	23.4	242/221	36	767	—
CG-Epi5	22	36.3 ± 0.8	0.39	2.9/4.5/9.4	1	47.6	—/370	38	804	0
Epi6	26	36.6 ± 0.9	0.88	51.4/15.3/54.7	1	111.1	854/594	47	705	—
CG-Epi6	26	36.3 ± 1.1	0.54	57.1/21.6/60.4	1	50.3	—/542	47	714	1.6
Bathy1	39	36.9 ± 0.9	1.2	54.3/23.4/60.4	1	94.7	1143/1007	46	807	—
CG-Bathy1	22	37.5 ± 1.8	1.04	60.0/26.1/64.2	1	211.7	—/988	46	792	4.8
Bathy2	37	64.6 ± 1.8	1.06	68.6/18.9/58.5	1	41.7	1023/654	40	771	—
CG-Bathy2	18	64.4 ± 1.5	0.77	68.6/21.6/58.5	1	130.1	—/704	41	816	2.4

Abbreviations: CDS, Coding DNA sequence; nr-CDS, non-redundant CDS.
[a](1) Raes *et al.* (2007); (2) Albertsen *et al.* (2013); (3) Narasingarao *et al.* (2012).
[b]nr-CDS: non-redundant CDS clustered at 80% similarity and 70% coverage.
[c]Parks *et al.*, 2014.

sequences against the NCBI-nr database, Pfam (Bateman *et al.*, 2004), arCOGS (Makarova *et al.*, 2015) and TIGRfams (Haft *et al.*, 2001) (cutoff *E*-value 10^{-5}). Based on sequence similarity against the non-redundant NCBI database, the best hit for each gene was determined and used to bin to top-level taxa. *Bona fide* Euryarchaeota genome fragments were defined as having >50% of the predicted open reading frames with best hits to other Euryarchaeota genes. The resulting sequences were used to screen for their presence in several metagenomes (in subsets of 20 million reads, where applicable): the TARA data sets (Sunagawa *et al.*, 2015), the GOS collection (Rusch *et al.*, 2007), the depth profiles collections from the subtropical gyres of North Atlantic (Bermuda Atlantic Time Series, BATS) and North Pacific (Hawaii Ocean Time-Series, HOT) (DeLong, 2006; Coleman and Chisholm, 2010), several Mediterranean Sea metagenomes at different depths (Ghai *et al.*, 2010; Quaiser *et al.*, 2011; Smedile *et al.*, 2012; Martin-Cuadrado *et al.*, 2015), and a number of deep ocean and cold waters metagenomes (Alonso-Saez *et al.*, 2012; Larsson *et al.*, 2014). The collections coming from the surroundings of hydrothermal vents published in Li *et al.* (2015) were also included. The screening was performed using Usearch6 (Edgar, 2010), with a cutoff of 95% identity over an alignment length of at least 50 bp (approximately species-level divergence, Konstantinidis and Tiedje, 2005). To compare the results among different data sets, the number of reads was normalized to the metagenome size and the sequence length. The final coverage results were expressed as the number of reads per kilobase of the fragment per gigabase of metagenome collection (rpkg). Only metagenomes in which any of the MG-III sequences recruited reads at over 3 rpkg, a total of 33 metagenomes, were used for genome assembly (Supplementary Table S1).

All the sequences obtained from these assemblies were binned together in order to cluster them by their tetranucleotide frequencies, GC content and coverage values (Supplementary Figure S2 and Supplementary Table S1). Tetranucleotide frequencies were computed using the 'wordfreq' program from the EMBOSS package (Rice *et al.*, 2000) and the coverage values were calculated as rpkg as described before. Only those clusters with >10 sequences and containing at least one gene marker with a clear affiliation to MG-III were retained. The phylogenetic assignment to MG-III was determined by the presence of at least one housekeeping gene in the same bin (see below). Following this method, a total of 375 genomic fragments >10 Kb could be classified into 10 different MG-III bins of sequences, Epi1, Epi2A, Epi2B, Epi2C, Epi3, Epi4, Epi5, Epi6, Bathy1 and Bathy2. We also considered 16 MG-III sequences that contained a ribosomal or a housekeeping gene but that could not be included in any of the bins by the criteria used (Supplementary Table S2).

In order to improve the completeness and remove the redundancy present in the initial MG-III bins,

Table 2 Environmental collections from where MG-III sequences were assembled

	Depth (m)	*Fraction size (μm)*	*Epi1*	*Epi3*	*Epi4*	*Epi5*	*Epi6*	*Epi2A*	*Epi2B*	*Epi2C*	*Bathy1*	*Bathy2*
Total, Kb			2950.4	707.0	631.3	259.7	848.7	542.7	564.7	305.0	1196.5	1061.4
ERR598993 (TARA_18)[a]	5	0.22–1.6	658.3									
ERR599073 (TARA_18)[a]	60	0.22–1.6	54.6									
ERR315859 (TARA_023)[a]	55	0.22–01.6					11.7					
ERR594297 (TARA_068)[a]	5	0.45–0.8	25.3									
ERR594294 (TARA_068)[a]	50	0.22–0.45	367.2				47.4					
ERR594348 (TARA_068)[a]	50	0.45–0.8	159.3									
ERR594335 (TARA_070)[a]	5	0.45–0.8	41.9									
ERR598942 (TARA_133)[a]	45	0.22–3		707.0		60.9						
ERR598983 (TARA_145)[a]	5	0.22–3				198.8			422.4	305.0		
ERR598996 (TARA_150)[a]	40	0.22–3	128.0									
ERR598976 (TARA_151)[a]	5	0.22–3	264.7									
ERR598986 (TARA_151)[a]	80	0.22–3	216.5									
MedDCM-OCT2007[b]	60	0.22–5	1034.7		34.5		733.8					
MedDCM-JUL2012[c]	75	0.22–5						542.7	142.3			
MedDCM-SEP2014[d]	60	0.22–5			596.7							
AD1000[e]	1000	0.22–5									38.7	
Med-Ae2–600mDeep[f]	600	0.22–5									1017.6	
Med-Io7–77mDCM[f]	77	0.22–5					55.8					
KM3[e]	3000	0.22–5									140.1	1059.6

[a]Sunaguawa *et al.* (2015). [b]Ghai *et al.* (2010). [c]Martin-Cuadrado *et al.* (2015). [d]This work. [e]Martin-Cuadrado *et al.* (2008). [f]Mizuno *et al.* (2016).

a second assembly was performed combining the sequences >10 Kb with the short paired-end Illumina reads of the metagenomes from where they were assembled (Tables 1 and 2 and Supplementary Figure S3). For each of the MG-III sequence bins, we used the BWA aligner (Li and Durbin, 2009; default parameters) to recover the short pair-reads that mapped onto the >10 Kb contigs. For each bin, these reads were then pooled and assembled together with the large DNA contigs previously assembled using SPAdes (Bankevich *et al.*, 2012). The final assemblies were termed 'composite genomes' (CGs), as they belong to similar MG-III cellular lineages (defined by the MG-III bins) but from different samples (Supplementary Table S3). The completeness of the reconstructed archaeal genomes was estimated by three different criteria and based on the presence of essential/core genes using HMMER (35, 112 and 53 genes (Raes *et al.*, 2007; Narasingarao *et al.*, 2012; Albertsen *et al.*, 2013)). An *E*-value $<10^{-5}$ and an alignment coverage >65% were used as cutoffs to define homologs of the essential/core genes. Analysis of the contamination within the CGs was performed using CheckM (Parks *et al.*, 2014) (Table 1). Average nucleotide identity (ANI) and conserved DNA fraction between reconstructed and/or reference genomes were calculated based on the whole-genome sequence as in Goris *et al.* (2007) (Supplementary Figure S4). GC content was calculated using the 'geecee' tool from the emboss package (Rice *et al.*, 2000).

Phylogenetic analysis

16S-rRNA and 23S-rRNA gene sequences detected in the MG-III genomic fragments were used to retrieve rRNA gene sequences from the most closely related euryarchaeal genomes and selected genome fragments in GenBank using BLAST (Altschul *et al.*, 1990). 16S-rRNA sequences from metagenome collections were screened and trimmed using ssu-align (Nawrocki, 2009). Archaeal 16S-rRNA and 23S-rRNA gene sequences were then aligned using MUSCLE (Edgar, 2004). Phylogenetic reconstructions were conducted by maximum likelihood using MEGA6-v.0.6 (Tamura-Nei model, 100 bootstraps, gamma distribution with (five discrete categories), all positions with <80% site coverage were eliminated) (Tamura *et al.*, 2013) (Supplementary Figure S5). For the protein trees of RecA, RpoB, SecY, geranylgeranylglyceryl phosphate synthase, DnaK, GyrA, GyrB, photolyase and rhodopsin (Supplementary Figures S6–S14), sequences were selected based on existing literature. Sequences were aligned using MUSCLE (Edgar, 2004) and a maximum likelihood tree was constructed using MEGA6-v.0.6 (Jones-Taylor-Thornton model, 100 bootstraps, gamma distribution with five discrete categories, positions with <80% site coverage were eliminated). Taxonomic affiliation of the selected bins was also determined by a phylogenomic tree based on concatenates of several ribosomal proteins (L13, S9, L5, S8, L6, S5, S12, S7, L11, L3, L4, L2, L22, S3, L14, S17, L15 and L18). A balanced taxonomic representation of other archaeal genomes was included as reference. Shared proteins were concatenated and aligned using Kalign (Lassmann and Sonnhammer, 2005) and a maximum likelihood tree was made using MEGA6-v.0.6.

Genome comparisons

Synteny among the CG-MGIII was examined with CIRCOS (Krzywinski *et al.*, 2009) and defined as arrays of contiguous genes in tracts of DNA >5 Kb and having >70% of identity. For each of the MG-III bins, non-redundant protein databases were

constructed clustering the coding DNA sequences with UCLUST (Edgar, 2010) (cutoff: 80% similarity in 70% of their length). These subsets of proteins were compared among themselves using a reciprocal best-hit analysis of putative homologs by BLASTP. Reciprocal relations were plotted using CYTOSCAPE (Shannon *et al.*, 2003). In order to identify the unique proteins of each of the bins, UCLUST was used with a cutoff of 30% similarity along 70% of their length.

Accession numbers

Mediterranean metagenomes used for recruitment are available at NCBI-BioProjects: PRJNA257723 (MedDCM-SEP2014, MedDCM-JUL2012 and MedDCM-OCT 2007), PRJNA305355 (Med-Io7–77mDCM, Med-Io16–70mDCM, Med-Io17–3500mDeep, Med-Ae1–75mDCM and Med-Ae2–600mDeep). Sequences >10 Kb and the reconstructed CGs genomes have been deposited in Bio-Project number: PRJNA335308. TARA metagenomes were downloaded from the European-Bioinformatics-Institute (http://www.ebi.ac.uk/services/tara-oceans-data).

Results and Discussion

General features of MG-III archaeal genomes

Following assembly and binning, we obtained 375 genomic fragments that clustered into 8 MG-III bins (Supplementary Figure S1). Six bins, Epi1–Epi6, were from epipelagic origin (photic zone) and contained a total of 386 genomic fragments with a total of 8.3 Mb. Two bins, Bathy1 and Bathy2, were from deep marine samples (aphotic zone) and contained 76 fragments for a total of 2.3 Mb. Manual inspection of the differential coverage of the sequences in each bin identified three subsets of Epi2, referred to as Epi2A, Epi2B and Epi2C. Further genomic comparisons indicated that these bins were very similar to each other (93–96% ANI, Supplementary Figure S4) and represent genomes from related species, likely within the same genus.

Remarkably, seven genome bins were formed by sequences primarily from a single sampling site (Table 2). The exception was Epi1, which includes sequences retrieved from nine different sites in the Mediterranean Sea, Atlantic and North-Pacific oceans. These findings suggest that the organisms represented by Epi1 are cosmopolitan in temperate epipelagic waters, whereas the other groups are only abundant enough to assemble from metagenomes at specific sites (endemic) or under transient environmental conditions causing significant growth (for example, blooms; see below).

To improve the analysis of each genome bin, a second assembly was performed and CGs were reconstructed using sequences from different samples and geographic origins (Supplementary Figure S1). These CGs are non-redundant and consist of genomic fragments from similar lineages of MG-III cells but not necessarily from the same sample. In this further assembly, subsets Epi2A, 2B and 2C were condensed into a single bin, CG-Epi2. Genomic features of the genome bins can be found in Tables 1 and 2 and the complete list of the MG-III contigs and the CGs are given in Supplementary Tables S2 and S3. Using the criteria of Narasingarao *et al.* (2012), the genome bins with highest degree of completeness were CG-Epi1 (85%), followed by CG-Epi2 (75%) and the mesopelagic CG-Bathy1 (64%). Based on the number of different variants of single copy genes in each bin, all our CGs contained a single microbial species each (Supplementary Table S4).

All MG-III bins had low GC content (36–36.8%) with the exception of Bathy2 (64.2%). Previously described MG-III sequences from different bathypelagic samples were all high GC (62.8%–65.4%) except for Guaymas32 (36.8%) (Li *et al.*, 2015). It has been noted that GC content tends to increase with depth (Romero *et al.*, 2009; Mizuno *et al.*, 2016). Selection for less nitrogen demand has been proposed as the main drive toward low genomic GC content in free-living marine bacterioplankton. In epipelagic waters, nitrogen is more likely to be the limiting nutrient, in contrast to the dark, energy-limited but relatively nitrogen-rich, deep ocean (Dufresne *et al.*, 2005; Swan *et al.*, 2013; Batut *et al.*, 2014; Giovannoni and Nemergut, 2014). Nevertheless, Bathy1 and Guaymas32 have similar low GC content to surface MG-III bins, suggesting that other factors might be also important.

In general, epipelagic MG-III bins were more genetically heterogeneous. Among the low GC-MGIII bins, the ANI varied from 68% to 85.4%, whereas the high GC-MGIII bins (Bathy2 is 90.8% similar to Cayman92) showed higher degrees of conservation, with ANIs ranging 89.5% to 96.2% (Supplementary Figure S4). This apparently higher diversity of the epipelagic groups may reflect the chemical and physical heterogeneity of surface water layers, which are submitted to stronger hydrodynamic, seasonal and geographical variations (Bryant *et al.*, 2015). In contrast, MG-III representatives from the deep ocean inhabit a more stable environment and might consequently be less diverse, with more homogenous genomes.

Phylogenetic affiliation of the genomic bins

Genes coding for rRNA are difficult to bin because (i) rRNA genes assemble poorly due to their conservation and duplication in genomes and (ii) they recruit metagenomic reads at much higher levels making coverage-based approaches impractical. Most of the rRNA sequences came from fosmid-libraries (Km3 and AD1000) and did not cluster within any of the bins described here. The only assigned 16S-rRNA sequence (372 bp) belonged to Bathy1 and it appears distantly related to the previously described OTU-D (Galand *et al.*, 2009) and DH148-W24 clusters (Lopez-Garcia *et al.*, 2001a,b) (Supplementary Figure S5a). A similar result was obtained with the 23S-rRNA gene identified in Bathy1 (Supplementary Figure S5b). Therefore, we

looked for other housekeeping genes that might be helpful to define the phylogenetic relationships of the novel MG-III with other archaea. We identified and constructed phylogenetic trees for RecA, RpoB, SecY, the geranylgeranylglyceryl phosphate synthase, DnaK and the two gyrase subunits, GyrA and GyrB (Supplementary Figures S6–S12). Although DnaK, GyrA and GyrB have a complex history of horizontal gene transfer (HGT) (Gribaldo *et al.*, 1999; Petitjean *et al.*, 2012; Raymann *et al.*, 2014), their phylogenetic analysis clearly showed the split between MG-II and MG-III sequences. The MG-III housekeeping genes retrieved from epipelagic waters clustered into two groups, one represented only by Epi2 and the other including Epi1, 3, 4, 5 and 6. Bathy2 appeared as a separate cluster from the epipelagic MG-III, and Bathy1 sequences appeared as the most divergent and basal branch. The phylogenomic analysis of the concatenated ribosomal proteins revealed a similar topology (Figure 1). The two epipelagic clusters shared similar GC content. Accordingly, they were named LowGC-MGIII (comprising two subclades: LowGC1-MGIII (Epi1, 3, 4 and 6) and LowGC2-MGIII (Epi2 and Guaymas32)), and a separate clade, containing bins exclusively of bathypelagic origin (Bathy2, Cayman92 and Guaymas31), was named HighGC-MGIII. Bin Epi5 lacks the ribosomal operon, but it was included into the LowGC1-MGIII based on the phylogenetic analysis of the other housekeeping genes (Supplementary Figures S11 and S12). Bathy1 consistently appeared as a separate basal branch, which might reflect the intermediate depth (600 m), location (Aegean Sea) and physicochemical conditions (highly saline, relatively warm and extremely oligotrophic) of the samples contributing sequences to this genomic bin. The position of Guaymas32 (retrieved from 1993 m), which clusters with Epi2 (5–75 m), might be explained by the presence of two different microbial species in the Guaymas32 bin (Li *et al.*, 2015). One appears to be most similar to the surface Epi2 sequences (80.8% ANI), while the other is closer to the deeper Bathy1 sequences (72.9% ANI) (also observed in the synteny plot of Figure 2a) (see below). Another plausible explanation is that Guaymas32 might be a surface

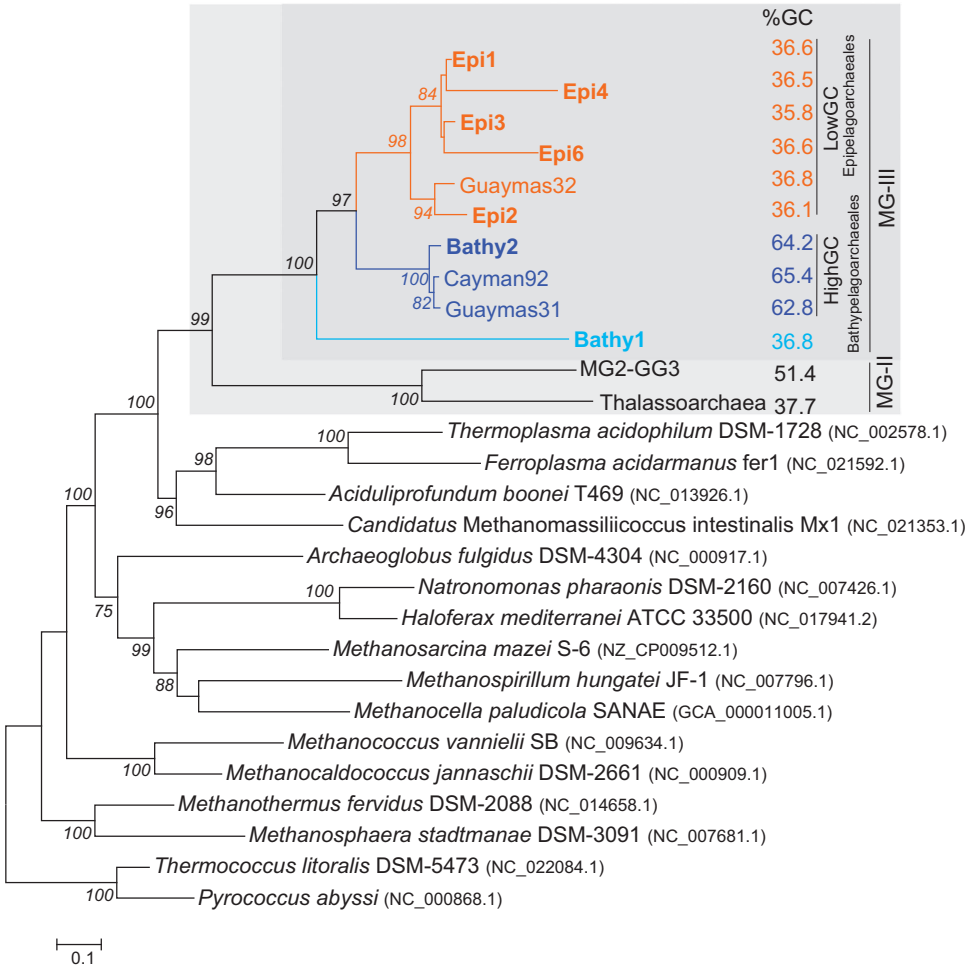


Figure 1 Maximum likelihood tree based on 18 ribosomal proteins concatenated present in draft MG-III archaeal genomes reconstructed from epipelagic and deep-sea metagenomes. Archaeal genomes from major orders of Euryarchaeota were included as references (accession number in brackets). Novel sequences from this work are shown in bold. Average GC content is shown on the right and colored depending on whether it is high or low GC. Only bootstrap values over >50% are shown.

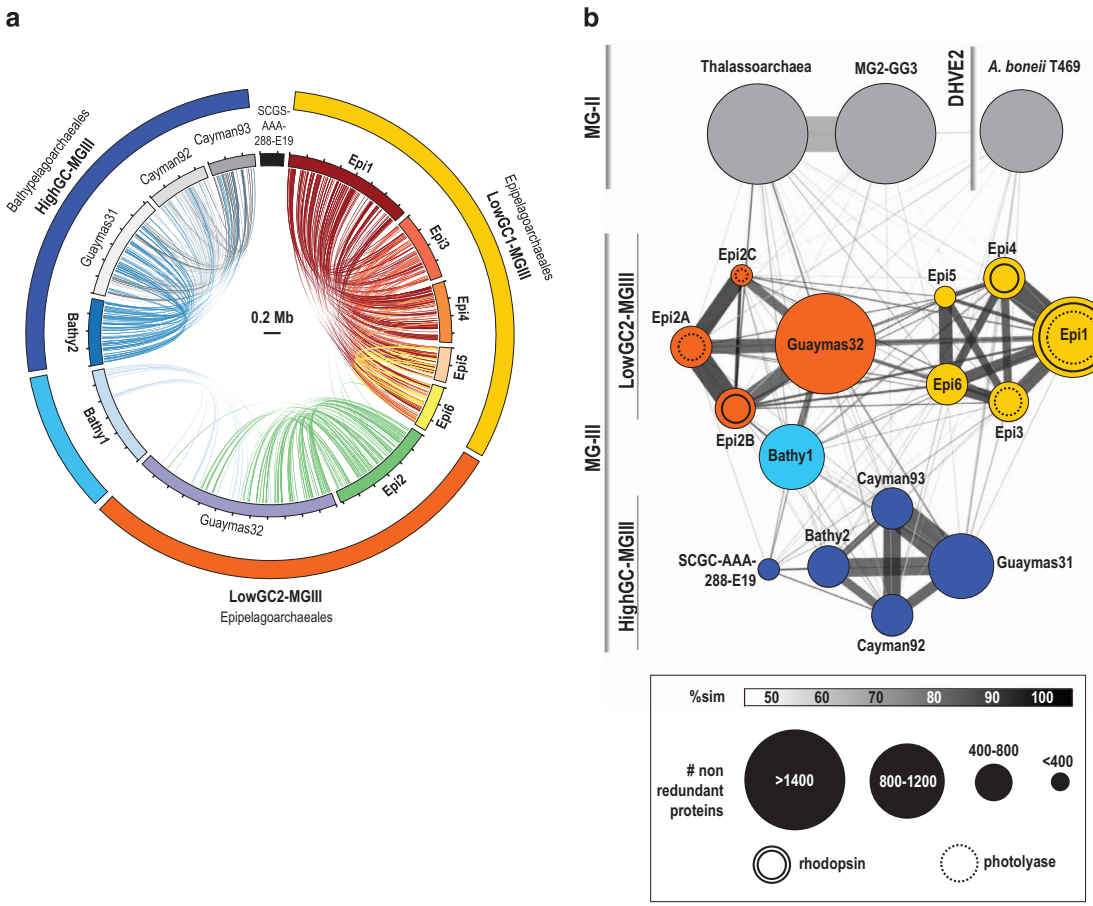


Figure 2 (**a**) Overview of genomic conserved synteny among the CG-MGIII genomes. Alignments >5 Kb over 70% identity are shown. A color code is used for each MG-III bin. (**b**) Amino-acid comparison among the MG-III bins. Sets of non-redundant proteins (cutoff of 80% similarity over 70% of their length) were compared through reciprocal BLASTP and the average amino-acid similarity was plotted. Each circle represents a genome bin. Circles are interconnected as a function of the percentage of shared proteins and colored in accordance with their similarity. Size of the bins and width of the lines are explained in the legend. Proteins of the MG-II MG2-GG3 (Iverson *et al*., 2012), Thalassoarchaea (Martin-Cuadrado *et al*., 2015) and the deep-sea hydrothermal vent Euryarchaeota (DHVE2) *Aciduliprofundum boneii* T469 were included in the analysis.

organism dragged to the bottom by the continuous flux of surface microbes and particles into the deep. Indeed, Guaymas sediments are surprisingly enriched in surface planktonic microbes (Edgcomb *et al*., 2002) when compared with other deep-sea sediments (Lopez-Garcia *et al*., 2003). However, the lack of rhodopsins and photolyases (discussed below), together with higher recruitments from deep data sets, would suggest that Guaymas32 is a *bona fide* deep inhabitant.

Synteny and gene content

To examine the conservation of synteny across the different genome bins, we performed an all-versus-all genome comparisons with the available sequences of MG-III (Figure 2a). Within the two groups of LowGC-MGIII bins, large fragments have the same genomic context while synteny blocks are not conserved between LowGC-MGIII and HigGC-MGIII. In the case of LowGC-MGIII, the highest synteny was found between Epi1 and Epi4 (54 block alignments, 62% of Epi4 genome size). For LowGC-MGIII, only Epi2 and Guaymas32 showed a significant synteny (56 block alignments, 38% of CG-Epi2). The low level of synteny between Bathy1 and other bins confirms that the microbes represented by this bin are very distant to the other LowGC-MGIII. Among the HighGC-MGIII bins, the highest synteny was found between Bathy2 and Guaymas31 (40 block alignments, 42% of CG-Bathy2) followed closely by Cayman92 and Guaymas31 (42 block alignments, 40.8% of the Cayman92 genome).

Non-redundant sets of proteins were obtained for each of the bins, including MG-II relatives, and compared between bins, retaining only the best hit for each protein and using a threshhold of 80% similarity. The relationships between bins were then plotted in the similarity network showed in Figure 2b. This protein content analysis supported the clustering observed in the phylogenomic tree (Figure 1). Bathy1 and SCGC-AAA-288-E19 appeared distantly associated with Guaymas32 and Guaymas31, respectively. MG-III bins Epi1 with Epi4 had the largest percentage of shared proteins (34.8%), followed by Epi2B and Guaymas32 (24%) and then Bathy2 and Guaymas31

(25%). Only 8% of Epi1 proteins were conserved in Epi2 and 0.5% in Bathy2. Although these numbers may be biased owing to the incomplete nature of the bins, they suggest that marine Euryarchaeota are very diverse and contain very different gene pools. Similar results were obtained by Deschamps *et al.* (2014) who found that the core genome of the MGII/III Euryarchaeota was only 15.6% of their pangenome, while their flexible genome was almost triple that of the Thaumarchaeota.

Metabolic functional inference

Several studies have suggested that marine Euryarchaeota have a significant role in the degradation of dissolved organic matter in marine waters, for example, dissolved amino acids (Ouverney and Fuhrman, 2000) or carbohydrates (Boutrif *et al.*, 2011). The presence of large peptidases related to protein degradation, together with enzymes for the use of fatty acids in the MG2-GG3 genome suggested that particles might be a habitat for MG-II Euryarchaeota (Iverson *et al.*, 2012; Orsi *et al.*, 2015). MG-II shared various features with the deep MG-III described by Li *et al.* (2015), suggesting that they might be aerobic heterotrophs that use proteins and polysaccharides as major energy source. In order to infer different lifestyles, the predicted open reading frames were functionally classified according the arCOG categories and their frequencies in the different genomes compared (Supplementary Tables S5 and S6 and Supplementary Figure S15).

Central carbon metabolism

MG-III genomes harbored enzymes for glycolysis, the tricarboxylic acid cycle and oxidative phosphorylation, indicating aerobic respiration (Supplementary Table S7). However, owing to the incomplete nature of these genomes, not all genes could be found, and some predictions need to be taken cautiously, especially for Bathy2. We found genes for the complete tricarboxylic acid cycle in LowGC-MGIII but three genes were absent in Bathy1. Remarkably, only the aconitase and the fumarase were found in Bathy2. As was observed in some MG-II (Martin-Cuadrado *et al.*, 2015), MG-III appears to possess most of the enzymes of the Embden–Meyerhof–Parnas (EMP) pathway for metabolism of hexose sugars, with the exception of the first and the last enzymes of the pathway. We were unable to find any other enzyme that could serve as an alternative for the missing glucokinase. For the final step of the EMP, we propose that phosoenolpyruvate synthase, found in all of our MG-III bins, might be able to function bi-directionally and substitute for the missing pyruvate kinase, allowing the EMP to function in both directions, gluconeogenic and glycolytic. Likewise, we found typical gluconeogenesis enzymes such as phosphoenolpyruvate carboxykinases in the LowGC-MGIII and Bathy1 bins, as well as subunits of the pyruvate/oxaloacetate carboxyltransferase in all the MG-III bins. We were unable to find glucose 1-dehydrogenase, gluconolactonase and 2-keto-3-deoxy gluconate aldolase homologs, suggesting that the Entner–Duodoroff hexose catabolic pathway is not present in the MG-III, unlike findings in other Euryarchaea (Makarova *et al.*, 1999; Makarova and Koonin, 2003; Hallam *et al.*, 2006).

Only a small number of amino-acid synthases were found in MG-III: cysteine in Bathy1 and Bathy2, glutamine in LowGC-MGIII, and for glutamate in all MG-III bins. Remarkably, many enzymes for *de novo* biosynthesis were missing, including those for synthesizing methionine, arginine, threonine, histidine, aromatic amino acids and branched amino acids (Supplementary Table S7). However, we observed multiple genes related with the uptake and transformation of peptides or amino acids in our MG-III bins, indicating that these organisms are capable of taking up amino acids from the environment and incorporating them into their proteins. For example, we found genes for permeases for lysine/arginine (all bins), histidine (Bathy2), glutamine (LowGC-MGIII and Bathy1), proline (LowGC-MGIII and Bathy1) and polar amino acids (Bathy2). Also, several ABC-transporter-systems were found for peptides and oligopeptides; for example, Dpp-ABC-type dipeptide/oligopeptide transporters (in all) and Liv-ABC-type branch amino-acid transporters (LowGC-MGIII and Bathy1). Several enzymes involved in the degradation of amino acids were also found, including dehydrogenases for alanine (all bins), glutamate (all bins), threonine (LowGC-MGIII and Bathy2) and proline (LowGC-MGIII), as well as several aminotransferases for branched-chain amino acids (LowGC-MGIII and Bathy1) and aspartate/tyrosine/aromatic aminotransferases (LowGC-MGIII and Bathy1). These findings suggest that there may be differences in the substrates used by the different MG-III groups. Indeed, although several subtilase-family proteases (arCOG00702 and arCOG02553) were present in all bins, some peptidases had limited distributions: dipeptidyl-aminopeptidases (LowGC-MGIII and Bathy1), C1A-peptidases (LowGC-MGIII), C25-peptidases (Bathy1), Xaa-Pro aminopeptidases (Bathy2), and several AprE-like subtilisins (arCOG06823, present in LowGC-MGIII and arCOG03610 present in Bathy1) (Supplementary Table S6).

Carbohydrates can be important carbon sources and, with the exception of Bathy1, several proteins with sugar-binding domains were found in all the bins (lectin and laminin-like). In the Epi6 bin, a cutin-like hydrolase was found (37% similar to a hydrolase from the Bacteriodetes *Rufibacter* sp. DG15C). Cutin is a polyester composed of hydroxyl/hydroxyepoxy fatty acids present in plants, and cutinases are produced by pathogenic fungi as extracellular degradative enzymes (Chen *et al.*, 1997). Lipo-oligosaccharide transport systems (*nodI/J*-like genes) and phosphonate transporters were found exclusively in the LowGC-MGIII. As observed in MG-II Thalassoarchaea (Martin-Cuadrado

et al., 2015), multidrug and antimicrobial peptide transporters (ABC-type) together with several permeases for drug/metabolites (RhaT-like family) were also abundant in all MG-III bins. Although the nature of the substrates is difficult to ascertain, these transporters may be involved in coping with high environmental concentrations of toxins such as those produced by cyanobacterial and algal blooms.

Oxygen. The presence of superoxide dismutase in all MG-III bins, together with several genes for alkyl-hydroperoxide reductases in LowGC1-MGIII and Bathy1, suggests that these microbes must cope with oxygen radicals. Complete cytochrome-*C* and B-B6 oxidase subunits operons were also found in LowGC1-MGIII and Bathy1 and Bathy2 bins. Copper-binding proteins and haloarchaeal-like halocyanins were found in proximity of these operons, an arrangement similar to that described for MG-II Thalassoarchaea (Martin-Cuadrado *et al*., 2015). It has been suggested that MG-II could be facultative anaerobes (Martin-Cuadrado *et al*., 2008; Belmar *et al*., 2011) and that sulfate could be used as terminal electron acceptor. Although no sulfate reductase-like proteins could be identified in our MG-III bins, several phosphate/sulfate permeases could be identified in Epi6 and Bathy2 and were also present in Guaymas31/32 and Cayman92. Pterin-based molybdenum enzymes (for example, sulfite oxidase, xanthine oxidase and dimethyl sulfoxide reductase) function under anaerobic conditions whereby their respective cofactors serve as terminal electron acceptors in respiratory metabolism (Schwarz *et al*., 2009). For Bathy2 (fosmid Km3–43-F08), a novel operon for the molybdopterin biosynthesis, was found (catalytic domains, MOCS1/S2/S3, have <55% similarities in the nr-database). However, we could not find any of the pterin-based enzymes.

Light-related genes. The presence of photolyases/cryptochromes among the LowGC-MGIII bins supports our hypothesis that they are *bona fide* epipelagic microbes (Figure 3a). Photolyases are proteins capable of photorepairing ultraviolet-induced pyrimidine dimers in the presence of light (Essen, 2006; Essen and Klar, 2006). Cryptochromes are proteins structurally similar to photolyases that act as blue light photoreceptors or regulators of the circadian rhythm (Cashmore *et al*., 1999) but that have lost the enzymatic photolyase activity (Chaves *et al*., 2011). Up to now, seven major classes of photolyase/cryptochrome families have been found (Scheerer *et al*., 2015). Interestingly, while the subunits found in Epi1 and Epi3 have similarity with eukaryotic cryptochromes (38–49%), the photolyases found in Epi2A and Epi2C bins have their highest similarities with Planctomycetales homologs (30–52%), suggesting potential inter-domain HGT events. Five related genes, a phytoene synthase, a phytoene-desaturase, an histidine kinase, a sugar-epimerase and one hypothetical protein, were found adjacent to the photolyase gene. At the equivalent genomic position, the aphotic Guaymas32 had neither the photolyase nor the associated genes mentioned above (downstream from a 23S-rRNA gene) (Figure 3a). The phylogenetic origin of the genes flanking the photolyases was analyzed and, in several cases, were most closely related to homologs from Bacteriodetes/Planctomycetes, again suggesting instances of HGT. These included a chaperone involved in protein secretion that was 76% similar to a *Rhodopirellula mairorica* homolog, a nitroreductase that was 75% similar to a *Gracilimonas tropica* homolog and a sugar-epimerase next to the photolyase that was 58% similar to a *Pirellula staleyi* protein. Likewise, a hypothetical protein adjacent to the photolyase in Epi1 and Epi3 was most closely related to eukaryotic genes, suggesting that this pair of genes may have been transferred together.

Epipelagic bins Epi1-2-3 all contained rhodopsins (Figure 2b) indicative of a photoheterotrophic lifestyle (Beja *et al*., 2000; Fuhrman *et al*., 2008; Inoue *et al*., 2013). In contrast, and consistent with previous reports (Deschamps *et al*., 2014; Li *et al*., 2015), Bathy1 and Bathy2 did not have rhodopsins. Phylogenetically, MG-III rhodopsins cluster with bacterial proteorhodopsins rather than with the euryarchaeal rhodopsins previously described for MG-II (Iverson *et al*., 2012; Martin-Cuadrado *et al*., 2014), suggesting that they may have been acquired by HGT from bacteria (Supplementary Figure S14). The analysis of key residues showed that all of these MG-III rhodopsins are proton pumps (Inoue *et al*., 2013) with a glutamine (Q) in the characteristic spectral tuning residue site indicating their ability to absorb light from the blue range (Supplementary Figure S16). In deeper waters (down to 300 m), only blue light remains available and blue rhodopsins are more suitable for generating energy. Therefore, epipelagic MG-III archaea seem to prefer low-light environments rather than the highly irradiated uppermost surface. Indeed, epipelagic MG-III bins recruited better from DCM or subsurface pelagic metagenomes (~50–70 m) than from surface (5 m) ones (see below). Genomic comparisons with MG-II rhodopsins (Martin-Cuadrado *et al*., 2014) revealed two new genomic contexts for this gene (Figure 3b). Interestingly, one of the clusters also contains one of the photolyase genes previously mentioned (Figure 3, contig Epi3-ERR598942-C530). Downstream from the rhodopsin genes, a gene for an unknown GYD domain protein was present. In cyanobacteria, proteins containing GYD and KaiC domains are involved in generating circadian rhythms (Chang *et al*., 2015). This raises the possibility that epipelagic MG-III Euryarchaeota may also have a circadian rhythm. A similar genome segment was found in two Guaymas32 sequences but, in these cases, the rhodopsin and the GYD domain-containing protein were absent.

The phylogenetic relationships of photolyases and rhodopsins, their proximity in at least one of the MG-III bins, together with the multiple putative HGT events observed in the nearby genes, leads us to

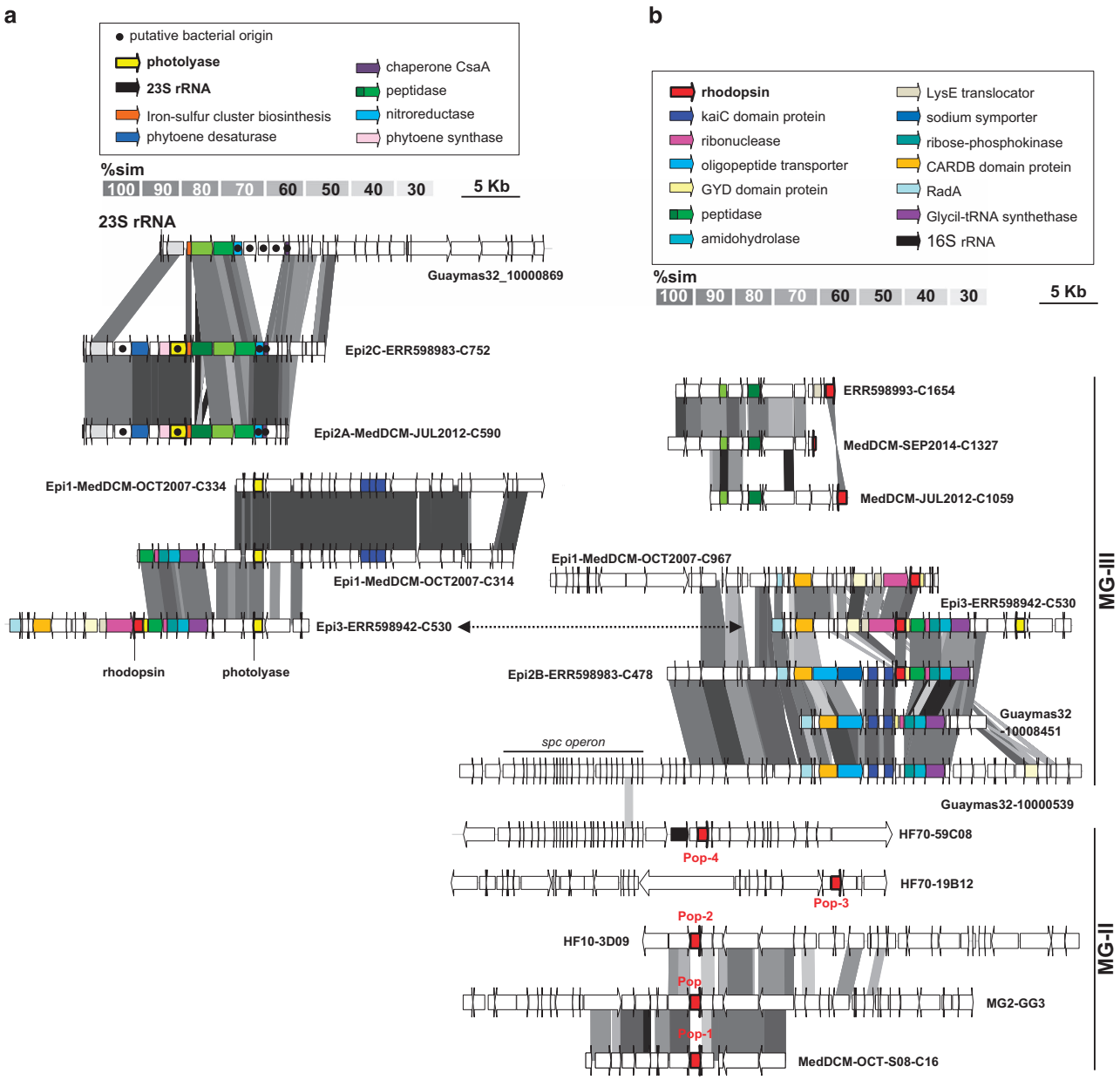


Figure 3 (**a**) Comparative genomic organization of MG-III sequences containing photolyases (in yellow). (**b**) Comparative genomic organization of MG-III sequences containing rhodopsins (in red) in context with other genomic fragments containing the MG-II Pop, Pop-1, Pop-2, Pop-3 and Pop-4 rhodopsins (bottom). Conserved genomic regions are indicated by gray shaded areas, gray intensity being a function of sequence similarity by TBLASTX. Particular open reading frames mentioned in the text are highlighted by a graphic code (see legend).

hypothesize an ancestral 'dark nature' for MG-III. These light-related genes would have been recently transferred from epipelagic bacteria to MG-III, probably long after the massive HGT events that have been detected prior to the diversification of several mesophilic archaeal clades, including MGII/III (Deschamps *et al.*, 2014; Lopez-Garcia *et al.*, 2016). The acquisition of proteorhodopsins, together with ultraviolet-protection photolyases, would have promoted a better adaptation to the oligotrophic surface waters allowing MG-III clades to expand into new photic niches.

Structural components

Cell envelope. One of the advantages of generating environmental fosmid sequences is that they allow the unequivocal assembly and detection of the so-called 'metagenomic islands' (Coleman *et al.*, 2006; Cuadros-Orellana *et al.*, 2007; Rodriguez-Valera *et al.*, 2009). These are clone-specific genome areas that, owing to their low coverage, are rarely assembled from metagenomic data sets but can be easily identified in reference-genome recruitment plots in the form of empty (or little populated) areas

with virtually no environmental homologs. One example can be observed in CG-Epi1. The area of the genome shown in Figure 4b (labeled with an asterisk) is enriched in genes needed for cell wall biosynthesis and contains several glycosyltransferases (type I/IV), together with polysaccharide

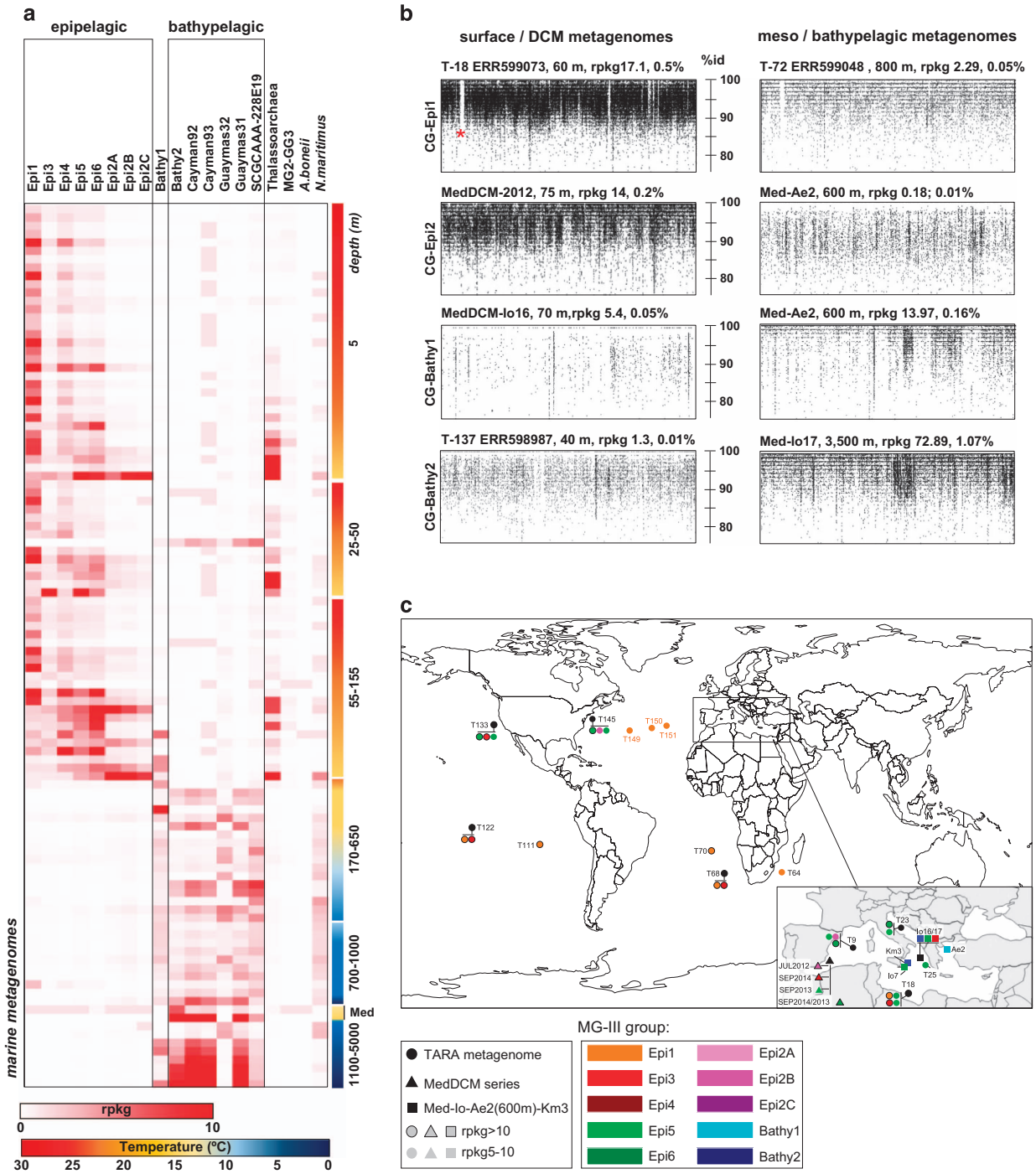


Figure 4 (**a**) Heat map of the number of rpkg of each CG-MGIII of this work together with the ones of Li *et al*. (2015), MG-II and other archaea genomes used as references in 106 different metagenomes from different geographical points and depths. Only those collections in which any of the MG-III sequences recruited rpkg > 1 were represented. (**b**) Recruitment plots of the CG-Epi1, CG-Epi2, CG-Bathy1 and CG-Bathy2 genomes in the metagenomes where they were better represented, from surface (< 200 m) and bathypelagic (> 500 m) (BLASTN-based, see Methods section). Rpkg and the percentages of the total of the reads with an identity bigger than > 95% are indicated. (**c**) Worldwide distribution of the CG-MGIII determined by metagenomic fragment recruitment against public metagenomic databases. Only samples where the CGs recruited rpkg > 5 are indicated in the map (cutoff: %identity > 95% in > 50 bp coverage). TARA spots are indicated by T#station.

synthases and genes for carbohydrate modification (acyltransferases and aminotransferases). The presence of several lipopolysaccharide biosynthesis proteins in all MG-III bins suggests a more complex cell envelope than a protein layer (S-layer). Adjacent to the CG-Epi1 island, we found a giant protein of 7258 amino acids with no similarity in sequence databases. These types of proteins have previously been observed in several bacterial and archaeal genomes (Reva and Tummler, 2008; Strom *et al.*, 2011) and have been hypothesized to have a role in defense against predation or in cell adhesion. Although we could not predict any function for it, the presence of lectin/glucanase domains (laminin_G3), glycosyl-transferase domains (RfaB), several beta-helix repeats and copper-binding domains (NosD) suggest an extracellular function. Large proteins (>5000 amino acids) with similar domains were also detected in other bins (Epi2-3-5). The similarity found between the giant proteins present in Guaymas31 and Bathy2 (90%) was remarkable.

Flagellum/Pili. Many archaeal surface structures are assembled by mechanisms related to the assembly of bacterial type IV pili (Lassak *et al.*, 2012). With the exception of Epi5, we found several sequences containing two concatenated *flaJ* genes (implicated in archaeal flagellum assembly) followed by a *flaI* gene (a transcriptional activator). Syntenic operons were also found among deep-MG-III in Li *et al.* (2015). However, these gene clusters are very different from the flagellar operon found in MG2-GG3 (Iverson *et al.*, 2012) or in any other Euryarchaeota described to date (Jarrell and McBride, 2008; Jarrell *et al.*, 2010). Although it has been claimed that the genes found might be enough to build a functional flagellum (Li *et al.*, 2015), the lack of a more complex gene cluster suggests that this operon might be involved in a secretion system translocating proteins rather than in cell motility.

Prevalence in the marine environment

To evaluate the relative abundance of the novel MG-III genomes, we used the non-redundant CGs to recruit reads from >200 metagenomic data sets that provide reasonably complete coverage of open-ocean waters from around the world. Among them, 106 gave values higher than one rpkg for any of the CGs tested (Figure 4 and Supplementary Table S1). Negative results are probably due to the small size of the data sets (for example, GOS) that may have poor representation of less abundant organisms. Although a considerable number of MG-III clones have been detected in cold waters such as the deep Atlantic layer of the central Arctic Ocean, (Galand *et al.*, 2009), the MG-III bins described here were not well represented in metagenomes from cold water regions such as polar regions (Alonso-Saez *et al.*, 2012), the Baltic (Larsson *et al.*, 2014) or the northeast subarctic Pacific (Allers *et al.*, 2013). This may suggest that there are other abundant MG-III groups present in high latitudes that have yet to be discovered. Even in warmer latitudes, our LowGC-MG-III bins only represent a small fraction of the total prokaryotic population of photic marine habitats. The highest abundance we found was for CG-Epi1 that accounted for 0.5% of the reads in the samples from the Mediterranean station TARA-018 (ERR599073 collection) (Figure 4b). The deep MG-III bins recruited slightly more. For instance, CG-Bathy2 recruited up to 1% of the reads in the deep sample Med-Io17 (3500 m).

Figure 4 shows a clear correlation of the two MG-III groups with depth (as already suggested by the origin of the assembled bins). Most LowGC-MGIII bins are only present in epipelagic collections, while the HighGC-MGIII plus the LowGC Bathy1 and Guaymas32 were clearly bathy or mesopelagic. CG-Epi1 seemed to be evenly distributed throughout the photic zone, but CG-Epi3, 5 and 6 increased at deeper waters (25–155 m, including the DCM) and the three CG-Epi2 showed an increase in even deeper photic zone waters. Bathy1 has its maximum at mesopelagic waters (Adriatic Sea 600 m), but it was also detected in colder bathypelagic waters (for example, the metagenomes from the Cayman-Rise and Guaymas Basin). CG-Bathy2 together with the Cayman and Guaymas bins revealed a strong correlation with deeper waters with much higher abundance in metagenomic collections <1000 m. These bins were more abundant in the warmer (13 °C) and saltier Mediterranean deep samples (KM3, 3000 m and Io17, 3500 m deep), although the temperature in most bathypelagic waters, where these microbes were detected (global ocean), typically decreases down to <5 °C. Overall, these numbers indicate that MG-III cells are relatively minor components of the archaeal communities in the photic and aphotic zones.

Using the Mediterranean DCM time series data sets, we found significant temporal variation in the abundance of the different GC bins despite a relatively constant abundance of reads attributable to euryarchaeal 16S rRNA genes (Supplementary Figure S17). For example, CG-Epi2A predominated in 2012, whereas CG-Epi6 was dominant in 2013 and CG-Epi4 in 2014. In the case of MG-II, it has been experimentally demonstrated that eukaryotic phytoplankton additions stimulate their growth in bottle incubations (Orsi *et al.*, 2015). Also, MG-II became one of the most abundant organisms (up to 40% of prokaryotes) in a phytoplankton bloom where diatoms, small flagellates and picophytoplankton dominated consecutively (Needham and Fuhrman, 2016). In order to know whether MG-II and the genomes of MG-III described here respond to similar blooming patterns, we measured the recruitment of available MG-II genomes in the metagenomes from which MG-III were assembled. The results show very low numbers for MG-II genomes in these samples, close to 100 times less than for MG-III genomes (Supplementary Figure S18). These data indicate that, despite being closely related and using similar substrates, MG-II and MG-III do not bloom concurrently.

Using published plankton-interactome data (Lima-Mendez *et al.*, 2015), we constructed an interaction network for MG-III archaea (Supplementary Figure S19). The results showed that MG-III coexists mainly with Metazoa and Dinophyta, which represented 50.6% and 23.5% of the total of interactions observed. These findings may indicate that MG-III cells could be attached to other organisms and only sporadically be released to the environment.

Conclusions

The photic zone of the oligotrophic ocean, one of the largest microbial habitats on Earth, has been extensively explored by molecular and genomic approaches (DeLong, 1992; DeLong *et al.*, 1999; Venter *et al.*, 2004; Rusch *et al.*, 2007; Sunagawa *et al.*, 2015). Nevertheless, many epipelagic microbes remain to be characterized. Using metagenomics, we have uncovered eight new groups of planktonic marine Euryarchaeota that likely represent novel taxonomic orders or at least families. Based on differences in genome content and sequence identity, we propose the following nomenclature: Epipelagoarchaeales for the LowGC-MGIII and Bathypelagoarchaeales for the HighGC-MGIII. A separate and basal clade with low GC content but apparently living in the dark ocean (Bathy1) has also been uncovered. Genome comparisons between these new groups together and previously described MG-III genomes (Li *et al.*, 2015) showed a marked differentiation between MG-III from photic and aphotic layers. Genomic analysis indicates that at least some representatives Epipelagoarchaeales (Epi1–Epi6) are planktonic photoheterotrophs. Two other groups with the Epipelagoarchaeales, Bathy1 and Guaymas32, lack genes indicating photoheterotrophy and are likely mesopelagic microbes with diverse metabolic capabilities. We hypothesize that the low GC content characteristic of the Epipelagoarchaeales may be an adaptation to the nitrogen limitation of surface waters. It is remarkable that all marine Euryarchaeota appear to possess similar metabolic profiles based on heterotrophic degradation of polymers and proteins (Iverson *et al.*, 2012; Martin-Cuadrado *et al.*, 2014; Li *et al.*, 2015; Orsi *et al.*, 2015). The broad diversity of marine microbes exploiting this habitat is likely a reflection of the enormous diversity of metabolic substrates available. Our data suggest a possible interaction of MG-III with eukaryotic cells and, more specifically, with metazoa.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

We are thankful to José de la Torre for editing the manuscript. We are thankful to L Gasperini and G Bortoluzzi of the Instituto di Geologia Marina (ISMAR), CNR, Bologna (Italy) for allowing PLG to participate in the Marmara2010 R/V Urania cruise during which part of the samples analyzed in this study were collected. This work was supported by projects MEDIMAX BFPU2013–48007-P from the Spanish Ministerio de Economía y Competitividad, MaCuMBA Project 311975 of the European Commission FP7, project AQUAMET II/2014/012 from the Generalitat Valenciana and by the French Agence Nationale de la Recherche (ANR-08-GENM-024–001,EVOL-DEEP). JHM was supported with a PhD fellowship from the Spanish Ministerio de Economía y Competitividad.

References

Albertsen M, Hugenholtz P, Skarshewski A, Nielsen KL, Tyson GW, Nielsen PH. (2013). Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nat Biotechnol* **31**: 533–538.

Alonso-Saez L, Waller AS, Mende DR, Bakker K, Farnelid H, Yager PL *et al.* (2012). Role for urea in nitrification by polar marine Archaea. *Proc Natl Acad Sci USA* **109**: 17989–17994.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. (1990). Basic local alignment search tool. *J Mol Biol* **215**: 403–410.

Allers E, Wright JJ, Konwar KM, Howes CG, Beneze E, Hallam SJ *et al.* (2013). Diversity and population structure of Marine Group A bacteria in the Northeast subarctic Pacific Ocean. *Isme J* **7**: 256–268.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS *et al.* (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19**: 455–477.

Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S *et al.* (2004). The Pfam protein families database. *Nucleic Acids Res* **32**(Dataabase issue): D138–D141.

Batut B, Knibbe C, Marais G, Daubin V. (2014). Reductive genome evolution at both ends of the bacterial population size spectrum. *Nat Rev Microbiol* **12**: 841–850.

Beja O, Suzuki MT, Koonin EV, Aravind L, Hadd A, Nguyen LP *et al.* (2000). Construction and analysis of bacterial artificial chromosome libraries from a marine microbial assemblage. *Environ Microbiol* **2**: 516–529.

Belmar L, Molina V, Ulloa O. (2011). Abundance and phylogenetic identity of archaeoplankton in the permanent oxygen minimum zone of the eastern tropical South Pacific. *FEMS Microbiol Ecol* **78**: 314–326.

Boutrif M, Garel M, Cottrell MT, Tamburini C. (2011). Assimilation of marine extracellular polymeric substances by deep-sea prokaryotes in the NW Mediterranean Sea. *Environ Microbiol Rep* **3**: 705–709.

Brochier-Armanet C, Boussau B, Gribaldo S, Forterre P. (2008). Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat Rev Microbiol* **6**: 245–252.

Bryant JA, Aylward FO, Eppley JM, Karl DM, Church MJ, DeLong EF. (2015). Wind and sunlight shape microbial diversity in surface waters of the North Pacific Subtropical Gyre. *ISME J* **10**: 1308–1322.

Cashmore AR, Jarillo JA, Wu YJ, Liu D. (1999). Cryptochromes: blue light receptors for plants and animals. *Science* **284**: 760–765.

Coleman ML, Chisholm SW. (2010). Ecosystem-specific selection pressures revealed through comparative population genomics. *Proc Natl Acad Sci USA* **107**: 18634–18639.

Coleman ML, Sullivan MB, Martiny AC, Steglich C, Barry K, Delong EF *et al.* (2006). Genomic islands and the ecology and evolution of Prochlorococcus. *Science* **311**: 1768–1770.

Cuadros-Orellana S, Martin-Cuadrado AB, Legault B, D'Auria G, Zhaxybayeva O, Papke RT *et al.* (2007). Genomic plasticity in prokaryotes: the case of the square haloarchaeon. *ISME J* **1**: 235–245.

Chang YG, Cohen SE, Phong C, Myers WK, Kim YI, Tseng R *et al.* (2015). Circadian rhythms. A protein fold switch joins the circadian oscillator to clock output in cyanobacteria. *Science* **349**: 324–328.

Chaves I, Pokorny R, Byrdin M, Hoang N, Ritz T, Brettel K *et al.* (2011). The cryptochromes: blue light photoreceptors in plants and animals. *Annu Rev Plant Biol* **62**: 335–364.

Chen S, Su L, Chen J, Wu J. (1997). Cutinase: characteristics, preparation, and application. *Biotechnol Adv* **31**: 1754–1767.

DeLong EF. (1992). Archaea in coastal marine environments. *Proc Natl Acad Sci USA* **89**: 5685–5689.

DeLong EF. (2006). Archaeal mysteries of the deep revealed. *Proc Natl Acad Sci U S A* **103**: 6417–6418.

DeLong EF, Preston CM, Mincer T, Rich V, Hallam SJ, Frigaard NU *et al.* (2006). Community genomics among stratified microbial assemblages in the ocean's interior. *Science* **311**: 496–503.

DeLong EF, Taylor LT, Marsh TL, Preston CM. (1999). Visualization and enumeration of marine planktonic archaea and bacteria by using polyribonucleotide probes and fluorescent in situ hybridization. *Appl Environ Microbiol* **65**: 5554–5563.

Deschamps P, Zivanovic Y, Moreira D, Rodriguez-Valera F, López-García P. (2014). Pangenome evidence for extensive inter-domain horizontal transfer affecting lineage-core and shell genes in uncultured planktonic Thaumarchaeota and Euryarchaeota. *Genome Biol Evol* **6**: 1549–1563.

Dufresne A, Garczarek L, Partensky F. (2005). Accelerated evolution associated with genome reduction in a free-living prokaryote. *Genome Biol* **6**: R14.

Edgar RC. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**: 113.

Edgar RC. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460–2461.

Edgcomb VP, Kysela DT, Teske A, de Vera Gomez A, Sogin ML. (2002). Benthic eukaryotic diversity in the Guaymas Basin hydrothermal vent environment. *Proc Natl Acad Sci USA* **99**: 7658–7662.

Essen LO. (2006). Photolyases and cryptochromes: common mechanisms of DNA repair and light-driven signaling? *Curr Opin Struct Biol* **16**: 51–59.

Essen LO, Klar T. (2006). Light-driven DNA repair by photolyases. *Cell Mol Life Sci* **63**: 1266–1277.

Fuhrman JA, Davis AA. (1997). Widespread archaea and novel bacteria from the deep sea as shown by 16S rRNA gene sequences. *March Ecol Prog Series* **150**: 275–285.

Fuhrman JA, McCallum K, Davis AA. (1992). Novel major archaebacterial group from marine plankton. *Nature* **356**: 148–149.

Fuhrman JA, Schwalbach MS, Stingl U. (2008). Proteorhodopsins: an array of physiological roles? *Nat Rev Microbiol* **6**: 488–494.

Galand PE, Casamayor EO, Kirchman DL, Potvin M, Lovejoy C. (2009). Unique archaeal assemblages in the Arctic Ocean unveiled by massively parallel tag sequencing. *ISME J* **3**: 860–869.

Galand PE, Gutiérrez-Provecho C, Massana R, Gasol JM, Casamayor EO. (2010). Inter-annual recurrence of archaeal assemblages in the coastal NW Mediterranean Sea (Blanes Bay Microbial Observatory). *Limnol Oceanogr.* **55**: 2117–2125.

Ghai R, Martin-Cuadrado A, Gonzaga A, Garcia-Heredia I, Cabrera R, Martin J *et al.* (2010). Metagenome of the Mediterranenan deep Chlorophyl Maximum stuidied by direct and fomid library 454 pyrosequencing. *ISME J* **4**: 1154–1166.

Ghai R, McMahon KD, Rodriguez-Valera F. (2012). Breaking a paradigm: cosmopolitan and abundant freshwater actinobacteria are low GC. *Environ Microbiol Rep* **4**: 29–35.

Giovannoni S, Nemergut D. (2014). Ecology. Microbes ride the current. *Science* **345**: 1246–1247.

Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* **57**: 81–91.

Gribaldo S, Lumia V, Creti R, Conway de Macario E, Sanangelantoni A, Cammarano P. (1999). Discontinuous occurrence of the hsp70 (dnaK) gene among Archaea and sequence features of HSP70 suggest a novel outlook on phylogenies inferred from this protein. *J Bacteriol* **181**: 434–443.

Haft DH, Loftus BJ, Richardson DL, Yang F, Eisen JA, Paulsen IT *et al.* (2001). TIGRFAMs: a protein family resource for the functional identification of proteins. *Nucleic Acids Res* **29**: 41–43.

Hallam SJ, Konstantinidis KT, Putnam N, Schleper C, Watanabe Y, Sugahara J *et al.* (2006). Genomic analysis of the uncultivated marine crenarchaeote Cenarchaeum symbiosum. *Proc Natl Acad Sci USA* **103**: 18296–18301.

Herndl GJ, Reinthaler T, Teira E, van Aken H, Veth C, Pernthaler A *et al.* (2005). Contribution of Archaea to total prokaryotic production in the deep Atlantic Ocean. *Appl Environ Microbiol* **71**: 2303–2309.

Huang Y, Gilna P, Li W. (2009). Identification of ribosomal RNA genes in metagenomic fragments. *Bioinformatics* **25**: 1338–1340.

Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**: 119.

Inoue K, Ono H, Abe-Yoshizumi R, Yoshizawa S, Ito H, Kogure K *et al.* (2013). A light-driven sodium ion pump in marine bacteria. *Nat Commun* **4**: 1678.

Iverson V, Morris RM, Frazar CD, Berthiaume CT, Morales RL, Armbrust EV. (2012). Untangling genomes from metagenomes: revealing an uncultured class of marine Euryarchaeota. *Science* **335**: 587–590.

Jarrell KF, Jones GM, Kandiba L, Nair DB, Eichler J. (2010). *S-layer glycoproteins and flagellins: reporters of archaeal posttranslational modifications*. Archaea 2010.

Jarrell KF, McBride MJ. (2008). The surprisingly diverse ways that prokaryotes move. *Nat Rev Microbiol* **6**: 466–476.

Karner MB, DeLong EF, Karl DM. (2001). Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* **409**: 507–510.

Konneke M, Bernhard AE, de la Torre JR, Walker CB, Waterbury JB, Stahl DA. (2005). Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* **437**: 543–546.

Konstantinidis KT, Tiedje JM. (2005). Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci USA* **102**: 2567–2572.

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D *et al.* (2009). Circos: an information aesthetic for comparative genomics. *Genome Res* **19**: 1639–1645.

Larsson J, Celepli N, Ininbergs K, Dupont CL, Yooseph S, Bergman B *et al.* (2014). Picocyanobacteria containing a novel pigment gene cluster dominate the brackish water Baltic Sea. *ISME J*.

Lassak K, Ghosh A, Albers SV. (2012). Diversity, assembly and regulation of archaeal type IV pili-like and non-type-IV pili-like surface structures. *Res Microbiol* **163**: 630–644.

Lassmann T, Sonnhammer EL. (2005). Kalign–an accurate and fast multiple sequence alignment algorithm. *BMC Bioinformatics* **6**: 298.

Li H, Durbin R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**: 1754–1760.

Li M, Baker BJ, Anantharaman K, Jain S, Breier JA, Dick GJ. (2015). Genomic and transcriptomic evidence for scavenging of diverse organic compounds by widespread deep-sea archaea. *Nat Commun* **6**: 8933.

Lima-Mendez G, Faust K, Henry N, Decelle J, Colin S, Carcillo F *et al.* (2015). Ocean plankton. Determinants of community structure in the global plankton interactome. *Science* **348**: 1262073.

Lopez-Garcia P, Lopez-Lopez A, Moreira D, Rodriguez-Valera F. (2001a). Diversity of free-living prokaryotes from a deep-sea site at the Antarctic Polar Front. *FEMS Microbiol Ecol* **36**: 193–202.

Lopez-Garcia P, Moreira D, Lopez-Lopez A, Rodriguez-Valera F. (2001b). A novel haloarchaeal-related lineage is widely distributed in deep oceanic regions. *Environ Microbiol* **3**: 72–78.

Lopez-Garcia P, Philippe H, Gail F, Moreira D. (2003). Autochthonous eukaryotic diversity in hydrothermal sediment and experimental microcolonizers at the Mid-Atlantic Ridge. *Proc Natl Acad Sci USA* **100**: 697–702.

Lopez-Garcia P, Zivanovic Y, Deschamps P, Moreira D. (2016). Bacterial gene import and mesophilic adaptation in archaea. *Nat Rev Microbiol* **13**: 447–456.

Makarova KS, Aravind L, Galperin MY, Grishin NV, Tatusov RL, Wolf YI *et al.* (1999). Comparative genomics of the Archaea (Euryarchaeota): evolution of conserved protein families, the stable core, and the variable shell. *Genome Res* **9**: 608–628.

Makarova KS, Koonin EV. (2003). Comparative genomics of Archaea: how much have we learned in six years, and what's next? *Genome Biol* **4**: 115.

Makarova KS, Wolf YI, Koonin EV. (2015). archaeal clusters of orthologous genes (arCOGs): an update and application for analysis of shared features between Thermococcales, Methanococcales, and Methanobacteriales*Life (Basel)* **5**: 818–840.

Martin-Cuadrado AB, Fontaine T, Esteban PF, del Dedo JE, de Medina-Redondo M, del Rey F *et al.* (2008). Characterization of the endo-beta-1,3-glucanase activity of *S. cerevisiae* Eng2 and other members of the GH81 family. *Fungal Genet Biol* **45**: 542–553.

Martin-Cuadrado AB, Garcia-Heredia I, Molto AG, Lopez-Ubeda R, Kimes N, Lopez-Garcia P *et al.* (2015). A new class of marine Euryarchaeota group II from the Mediterranean deep chlorophyll maximum. *ISME J* **9**: 1619–1634.

Martin-Cuadrado AB, Pasic L, Rodriguez-Valera F. (2014). Diversity of the cell-wall associated genomic island of the archaeon *Haloquadratum walsbyi*. *BMC Genomics* **16**: 603.

Martin-Cuadrado AB, Rodriguez-Valera F, Moreira D, Alba JC, Ivars-Martinez E, Henn MR *et al.* (2008). Hindsight in the relative abundance, metabolic potential and genome dynamics of uncultivated marine archaea from comparative metagenomic analyses of bathypelagic plankton of different oceanic regions. *ISME J* **2**: 865–886.

Massana R, DeLong EF, Pedros-Alio C. (2000). A few cosmopolitan phylotypes dominate planktonic archaeal assemblages in widely different oceanic provinces. *Appl Environ Microbiol* **66**: 1777–1787.

Mizuno CM, Ghai R, Saghai A, Lopez-Garcia P, Rodriguez-Valera F. (2016). Genomes of abundant and widespread viruses from the deep ocean. *MBio* **7**: 4.

Narasingarao P, Podell S, Ugalde JA, Brochier-Armanet C, Emerson JB, Brocks JJ *et al.* (2012). De novo metagenomic assembly reveals abundant novel major lineage of Archaea in hypersaline microbial communities. *ISME J* **6**: 81–93.

Nawrocki EP. (2009). *Structural RNA homology search and alignment using covariance models*, PhD thesis. Washington University in Saint Louis, School of Medicine, St. Louis, MO, USA.

Needham DM, Fuhrman JA. (2016). Pronounced daily succession of phytoplankton, archaea and bacteria following a spring bloom. *Nat Microbiol*.

Orsi WD, Smith JM, Wilcox HM, Swalwell JE, Carini P, Worden AZ *et al.* (2015). Ecophysiology of uncultivated marine euryarchaea is linked to particulate organic matter. *ISME J* **9**: 1747–1763.

Ouverney CC, Fuhrman JA. (2000). Marine planktonic archaea take up amino acids. *Appl Environ Microbiol* **66**: 4829–4833.

Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. (2014). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**: 1043–1055.

Peng Y, Leung HC, Yiu SM, Chin FY. (2012). IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* **28**: 1420–1428.

Petitjean C, Moreira D, Lopez-Garcia P, Brochier-Armanet C. (2012). Horizontal gene transfer of a chloroplast DnaJ-Fer protein to Thaumarchaeota and the evolutionary history of the DnaK chaperone system in Archaea. *BMC Evol Biol* **12**: 226.

Qin W, Amin SA, Martens-Habbena W, Walker CB, Urakawa H, Devol AH *et al.* (2014). Marine ammonia-oxidizing archaeal isolates display obligate mixotrophy and wide ecotypic variation. *Proc Natl Acad Sci USA* **111**: 12504–12509.

Quaiser A, Zivanovic Y, Moreira D, Lopez-Garcia P. (2011). Comparative metagenomics of bathypelagic plankton and bottom sediment from the Sea of Marmara. *ISME J* **5**: 285–304.

Raes J, Korbel JO, Lercher MJ, von Mering C, Bork P. (2007). Prediction of effective genome size in metagenomic samples. *Genome Biol* **8**: R10.

Raymann K, Forterre P, Brochier-Armanet C, Gribaldo S. (2014). Global phylogenomic analysis disentangles the complex evolutionary history of DNA replication in archaea. *Genome Biol Evol* **6**: 192–212.

Reva O, Tummler B. (2008). Think big–giant genes in bacteria. *Environ Microbiol* **10**: 768–777.

Rice P, Longden I, Bleasby A. (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* **16**: 276–277.

Rodriguez-Valera F, Martin-Cuadrado AB, Rodriguez-Brito B *et al.* (2009). Explaining microbial population genomics through phage predation. *Nat Rev Microbiol* **7**: 828–836.

Romero H, Pereira E, Naya H, Musto H. (2009). Oxygen and guanine-cytosine profiles in marine environments. *J Mol Evol* **69**: 203–206.

Rusch DB, Halpern AL, Sutton G, Heidelberg KB, Williamson S, Yooseph S *et al.* (2007). The Sorcerer II Global Ocean Sampling Expedition: Northwest Atlantic through Eastern Tropical Pacific. *PLoS Biol* **5**: e77.

Scheerer P, Zhang F, Kalms J, von Stetten D, Krauss N, Oberpichler I *et al.* (2015). The class III cyclobutane pyrimidine dimer photolyase structure reveals a new antenna chromophore binding site and alternative photoreduction pathways. *J Biol Chem* **290**: 11504–11514.

Schwarz G, Mendel RR, Ribbe MW. (2009). Molybdenum cofactors, enzymes and pathways. *Nature* **460**: 839–847.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D *et al.* (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**: 2498–2504.

Smedile F, Messina E, La Cono V, Tsoy O, Monticelli LS, Borghini M *et al.* (2012). Metagenomic analysis of hadopelagic microbial assemblages thriving at the deepest part of Mediterranean Sea, Matapan-Vavilov Deep. *Environ Microbiol* **3**: 1462–2920.

Strom SL, Brahamsha B, Fredrickson KA, Apple JK, Rodriguez AG. (2011). A giant cell surface protein in Synechococcus WH8102 inhibits feeding by a dinoflagellate predator. *Environ Microbiol*.

Sunagawa S, Coelho LP, Chaffron S, Kultima JR, Labadie K, Salazar G *et al.* (2015). Ocean plankton. Structure and function of the global ocean microbiome. *Science* **348**: 1261359.

Swan BK, Chaffin MD, Martinez-Garcia M, Morrison HG, Field EK, Poulton NJ *et al.* (2014). Genomic and metabolic diversity of Marine Group I Thaumarchaeota in the mesopelagic of two subtropical gyres. *PLoS One* **9**: e95380.

Swan BK, Tupper B, Sczyrba A, Lauro FM, Martinez-Garcia M, Gonzalez JM *et al.* (2013). Prevalent genome streamlining and latitudinal divergence of planktonic bacteria in the surface ocean. *Proc Natl Acad Sci USA* **110**: 11463–11468.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* **30**: 2725–2729.

Tyson GW, Chapman J, Hugenholtz P, Allen EE, Ram RJ, Richardson PM *et al.* (2004). Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* **428**: 37–43.

Vavourakis CD, Ghai R, Rodriguez-Valera F, Sorokin DY, Tringe SG, Hugenholtz P *et al.* (2016). Metagenomic insights into the uncultured diversity and physiology of microbes in four hypersaline Soda Lake Brines. *Front Microbiol* **7**: 211.

Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA *et al.* (2004). *Environmental genome shotgun sequencing of the Sargasso Sea. Science* **304**: 66–74.

Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)

SUPPLEMENTARY INFORMATION

Supplementary Figure S1. Schematic representation of the assembly and binning procedure.

Supplementary Figure S2. Binning of the MG-III genomic fragments. a. Principal component analysis of tetranucleotide frequencies of the MG-III DNA fragments. Reference sequences are shown as larger circles: *Nitrosopumilus maritimus SCM1*, *Aciduliprofundum boneii* T469, MGII-GG3, MGII-Thalassoarchaea, the single amplified genome (SAG) SCGC-AAA288-E19 and the DNA fragments from Cayman92, Cayman93, Guaymas31 and Guaymas32. **b.** Heat-map of the number of reads per kilobase per gigabase of metagenome collection (rpkg) of the MG-III sequences recruiting in 33 different metagenomes (only those in which any of the MG-III sequences recruited over rpkg>3 were considered for the binning analysis).

Supplementary Figure S3. Distribution of the assembled MG-III sequences among the metagenomes of the Mediterranean series and TARA collections ordered by depth. Size fraction is indicated above each column.

Supplementary Figure S4. Average nucleotide identity (ANI) among the MG-III genome bins (in bold the sequences published in this work). Dendrogram showing the similarity among the bins is shown in the *y* axis.

Supplementary Figure S5. 16S and 23S-rRNA phylogeny. a. Maximum-likelihood 16S-rRNA gene tree showing the relationship of the MG-III with other archaea. Circles at nodes in major branches indicate bootstrap support (see legend). Scale bar represents the estimated number of substitutions per site. The different MG-III clusters are indicated by different colours and named following Galand *et al.* 2008 and this work. Sampling location is indicated in each of the sequences. Those sequences from samples from more than 500m deep are underlined. **b.** Maximum likelihood 23S-rRNA

gene tree showing the relationship of the MG-III with other archaea. (Picture features as explained before).

Supplementary Figures S6-S12. Maximum likelihood phylogenetic trees for the housekeeping genes RecA, RpoB, SecY, geranylgeranylglyceryl phosphate synthase, DnaK, GyrA and GyrB. Protein sequences from this study are indicated in bold and coloured accordingly with the sequence bin (see legend in Supplementary Figure S4). Genomic bins of MG-II and MG-IIII groups are indicated. Bootstrap values over 50 are indicated.

Supplementary Figure S13. Maximum likelihood phylogenetic tree of the photolyases and cryptochromes found among the MG-III bins. Protein sequences from this study are indicated in bold and coloured accordingly with their kingdom affiliation (see legend). Bootstrap values over 50 are indicated.

Supplementary Figure S14. Maximum likelihood phylogenetic tree showing the relationship of the MG-III rhodopsins with other bacterial and archaeal rhodopsins. Protein sequences from this study are indicated in bold and coloured accordingly with the sequence bin (see legend in Supplementary Figure S4). Following the nomenclature of Iverson et al. (2013), Clade A and Clade B of rhodopsins is shown. In blue are marked Pop, Pop-1, Pop-2, Pop-3 and Pop-4 euryarchaeal rhodopsins previously described. Numbers at nodes in major branches indicate bootstrap support (shown as percentages and only those >50%). Scale bar represents the estimated number of substitutions per site.

Supplementary Figure S15. Distribution of arCOG functional classes. Percentage of arCOGs predicted in the MGIII bins described in this work and MG-II marine euryarchaeal genomes MG2-GG3 and Thalasoarchaea. All genes (**a**) and genes found only in one of the MGIII bins (**b**) are indicated. Asterisks indicate categories where a significant variation was found comparing the epipelagic and pelagic MGIII.

Supplementary Figure S16. Alignment of the MG-III rhodopsins with other cloned rhodopsins sequences. Identical residues are indicated in red. Residues in blue are conserved in more than 70% of the sequences. Key amino acids for rhodopsins functionality (listed herein with *G. pallidula* numbering) are marked by colours: Lys336 (K) binds retinal, and Asp164 (D) and Glu175 (E) function as Schiff base proton acceptor and donor, respectively. Glutamine (Q) in position 172 (*) in the MG-III rhodopsins sequences indicates an absorption maxima at the blue spectrum range. Letters (G) and (B) in the name of the sequences indicate the range of the spectrum. (The GenBank accession numbers of the sequences used for the alignment are as follows: Pop-2 HF10_3D09, 82548293; Pop-3 HF70_19B12, 82548286; Pop-4 HF70_59C08, 77024964; eBAC49C08, AAY82659; HF130_81H07, 119713419; HF10_49E08, 119713779; eBAC20E09, AAS73014; HOT75m4, AAK30179; eBAC31A08 (SAR86), AAG10475; SAR86E, WP_008490645; *C.* Pelagibacter ubique HTCC1062 (SAR11), YP_266049; *Pelagibacter* sp. HTCC7211, WP_008544914; gammaproteobacteria HTCC2207 (SAR92), EAS48197; *G. pallidula*, WP_006008821; *Dokdonia donghaensis* MED134, ZP_01049273; MedDCM-JUL2012-C3793, KP211865; MedDCM-OCT2007-C1678, KP211832; MedDCM-OCT-S08-C16, ADD93192; *Exiguobacterium sibiricum* 255-15, ACB60885; *Exiguobacterium* sp. AT1b, WP_012726785; *Haloarcula marismortui* ATCC 43049, YP_136594.

Supplementary Figure S17. Classification of the DCM-metagenomes reads using the RDP (16S-rRNA) database. Only those genera which represented more than 1% were represented.

Supplementary Figure S18. Number of reads per kilobase per gigabase of metagenome collection (rpkg) for the composite genome CG-Epi1 MG-III (ordered from minus to major) compared with those obtained for the MG-II (MG2-GG3 (Iverson *et al.* 2013) and the Thalassoarchaea (Martin-Cuadrado *et al.* 2014)). Other CG MG-III were also included in the graph.

Supplementary Figure S19. Interactions (showed as percentages) calculated by Mendez-Lima *et al.,* (2015) for the MG-III with other organisms. Data extracted from Supplementary table W7 in Mendez-Lima *et al.,* (2015).

SUPPLEMENTARY TABLES

Supplementary Table S1. List of metagenomes used for recruitments in Figure 4a and Supplementary Figure 2b, sorted by temperature and depth.

Supplementary Table S2. List of sequences of each of the MG-III sequence bins.

Supplementary Table S3. List of sequences of the MG-III composite genomes (CG-MGIII).

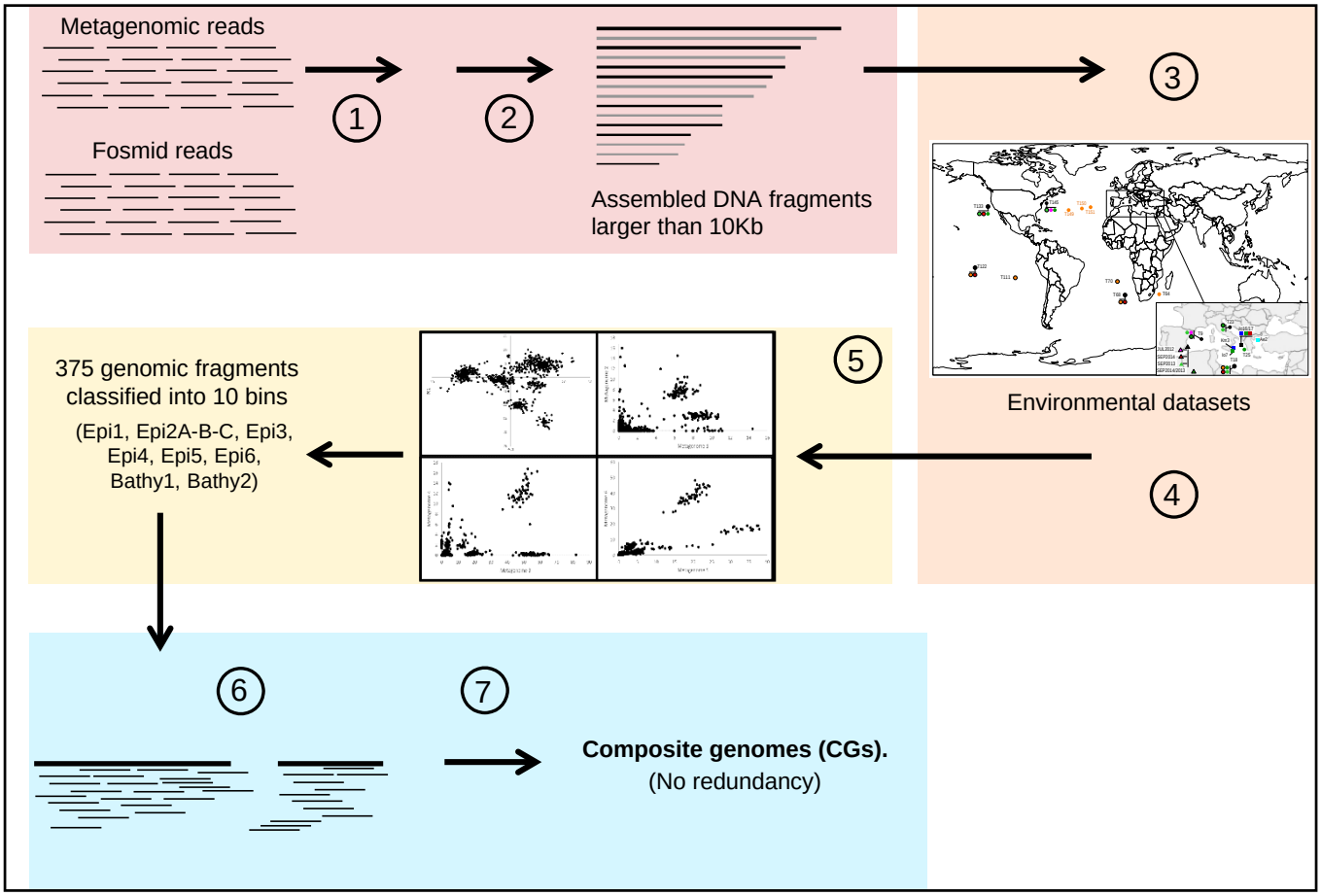
Supplementary Table S4. Housekeeping genes found in the MG-III bins and the CG-MG-III bins (as Narasingarao *et al.*, (2012).

Supplementary Table S5. CG-MGIII-protein categories based on the arCOG database.

Supplementary Table S6. Classification of the CG-MGIII unique CDSs based on the arCOG classification.

Supplementary Table S7. List of genes involved in MG-III metabolic pathways.

Supplementary Figure S1



Assembly and annotation of euryarchaeal sequences

1 **Assembly.** Metagenomic reads from MedDCM-JUL2012, MedDCM-SEP2014, Med-Io7-77mDCM, Med-Ae2-600mDeep and the fosmid reads from MedDCM-OCT2007, KM3 and AD1000 were independent and systematically assembled using IDBA_UD.

2 **Annotation.** Assembled DNA fragments were annotated using the NCBI nr-database, Pfam, COGs, arCOGs and TIGRfam. Only those fragments which have more than 50% of their ORFs annotated as Euryarchaeota were taken for further analysis (in black).

Retrieving of euryarchaeal sequences

3 **Recruitment.** Search of Euryarchaeota sequences within different marine databases. Only those metagenomes where our initial genomic fragments recruited more than 3 rpkg were selected.

4 **Assembly and annotation** of the selected metagenomes (as described previously). DNA fragments classified as euryarchaeota were collected.

Binning of MG-III by tetranucleotide frequencies, %GC and differential coverage

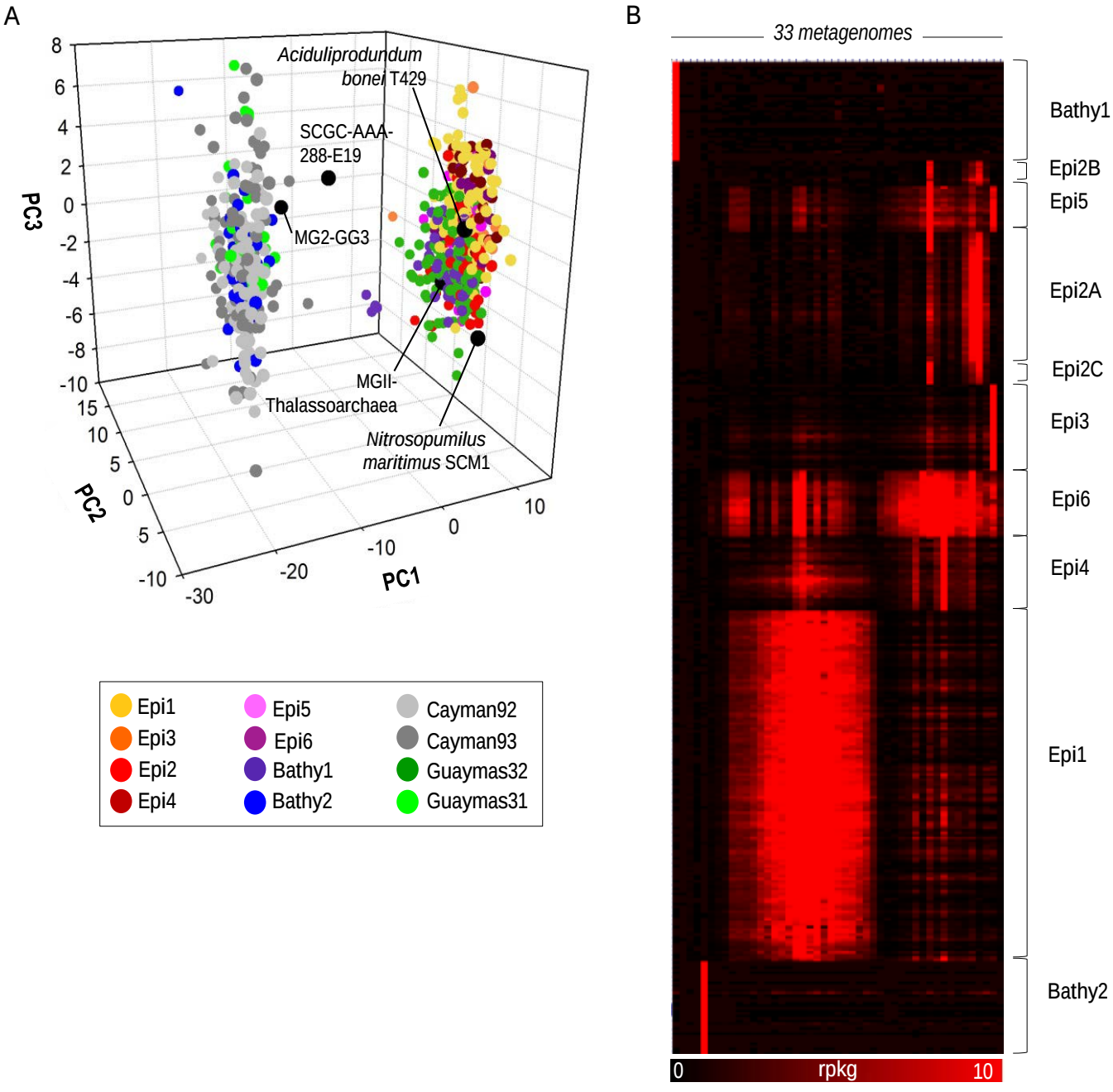
5 **Binning** of the MG-III sequences by tetranucleotide frequencies, %GC and differential coverage among 33 metagenomes.

Re-assembly

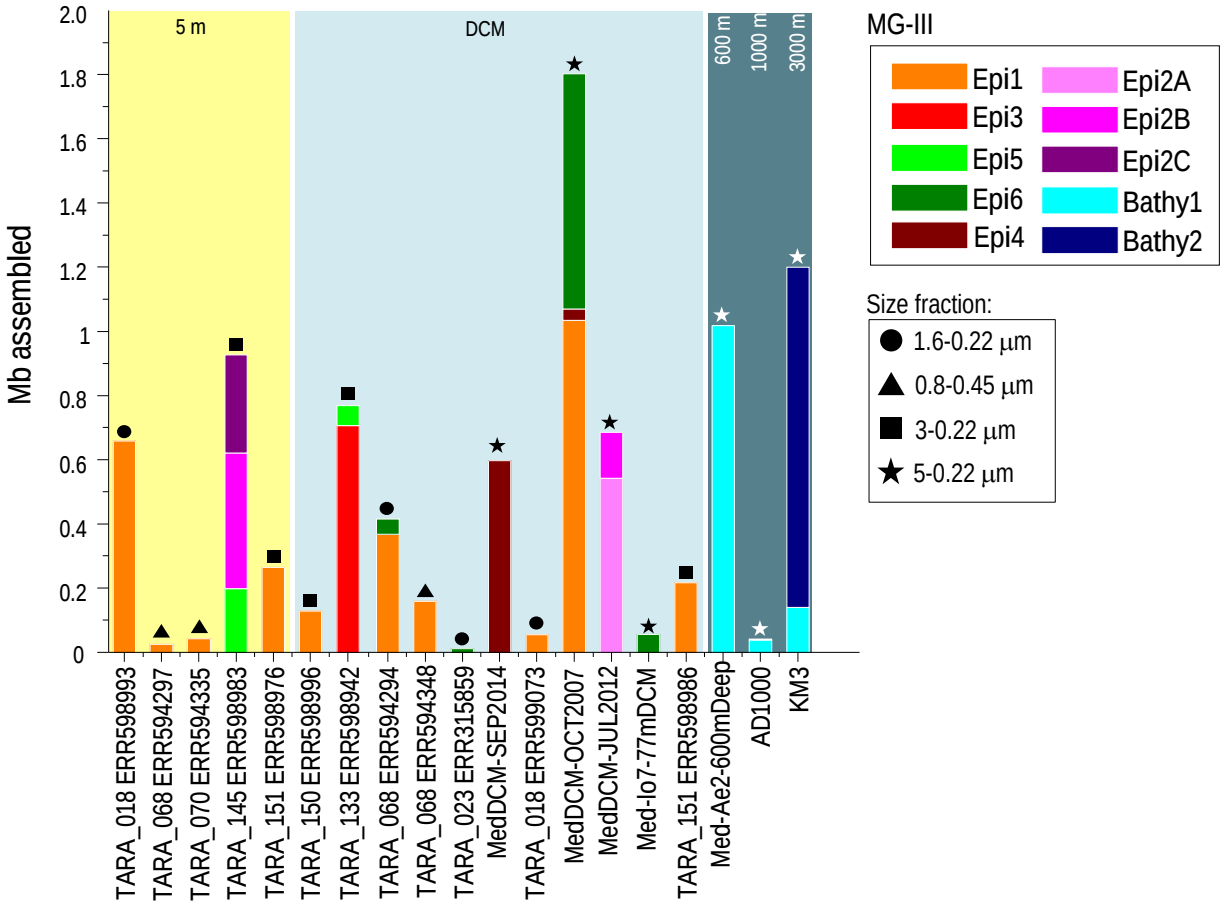
6 **Aligment** of the sequences >10 Kb of one specific bin against the metagenomic reads where those contigs came from using BWA. The process was made for all the bins. The reads extracted were used in the following step.

7 **Assembly** of the reads with SPADES using the DNA fragments for scaffolding with the parameter –trusted_contigs. The re-assembly was done for all the bins and their products are called «composite genome» (CG).

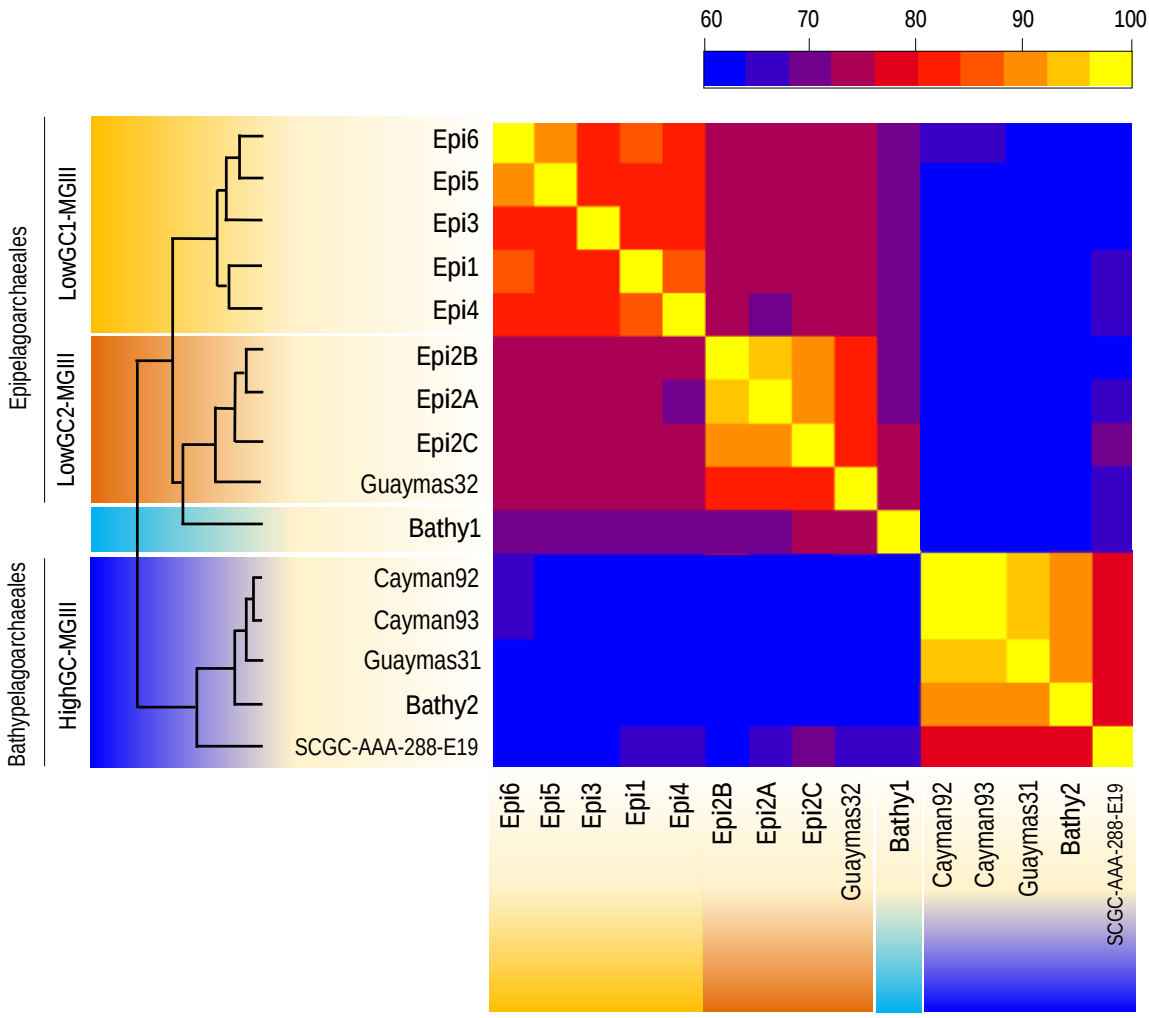
Supplementary Figure S2



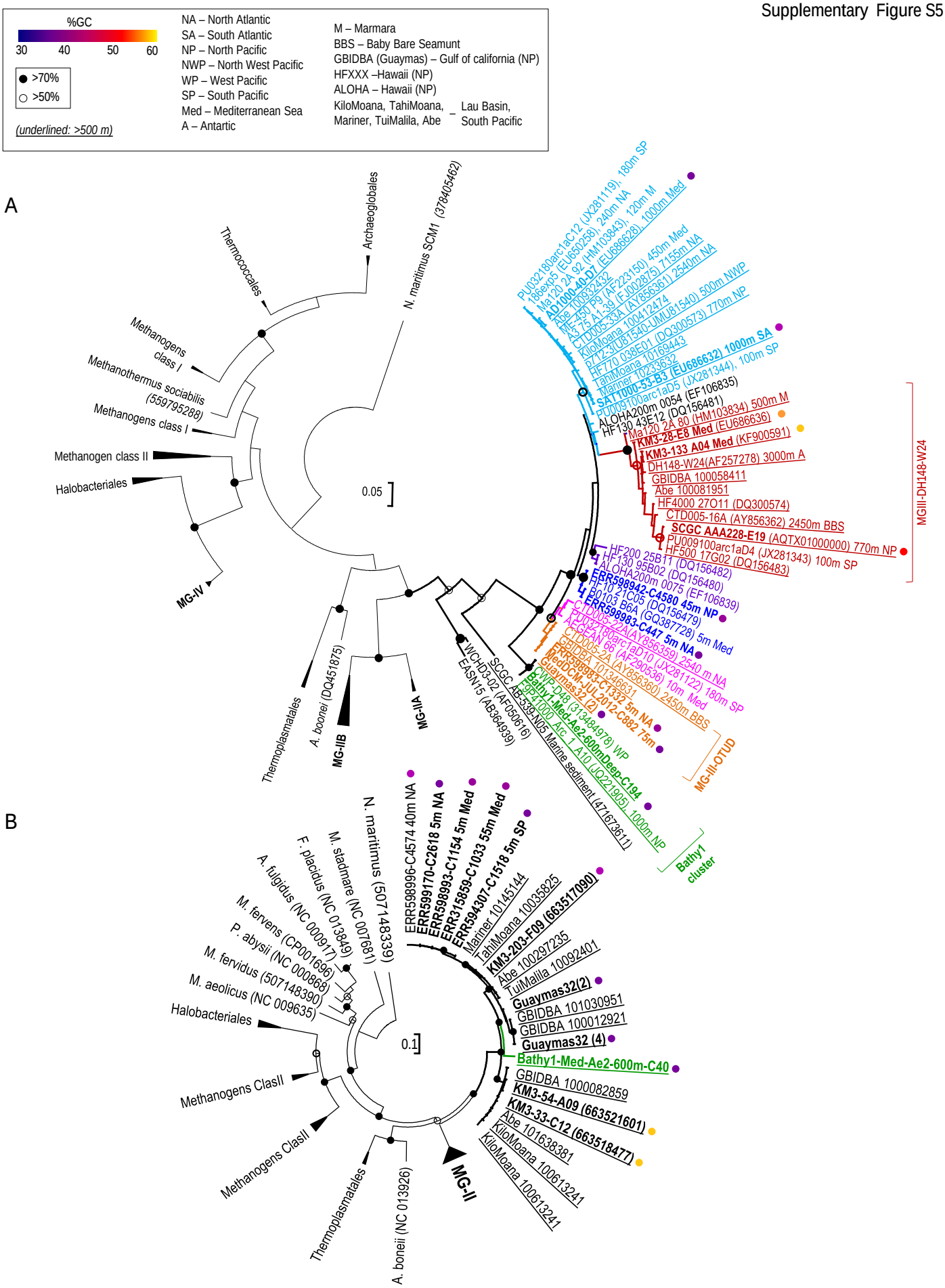
Supplementary Figure S3



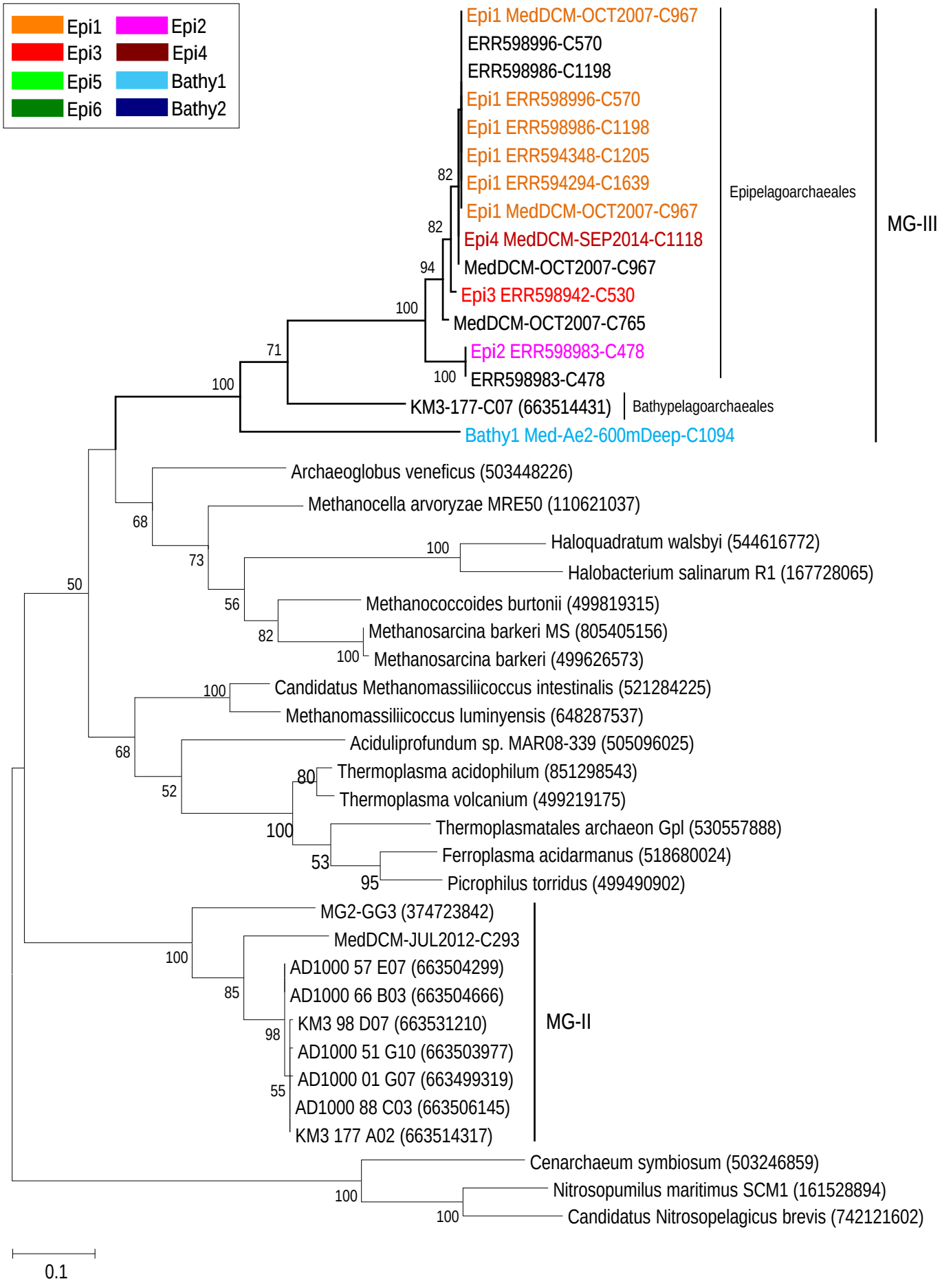
Supplementary Figure S4



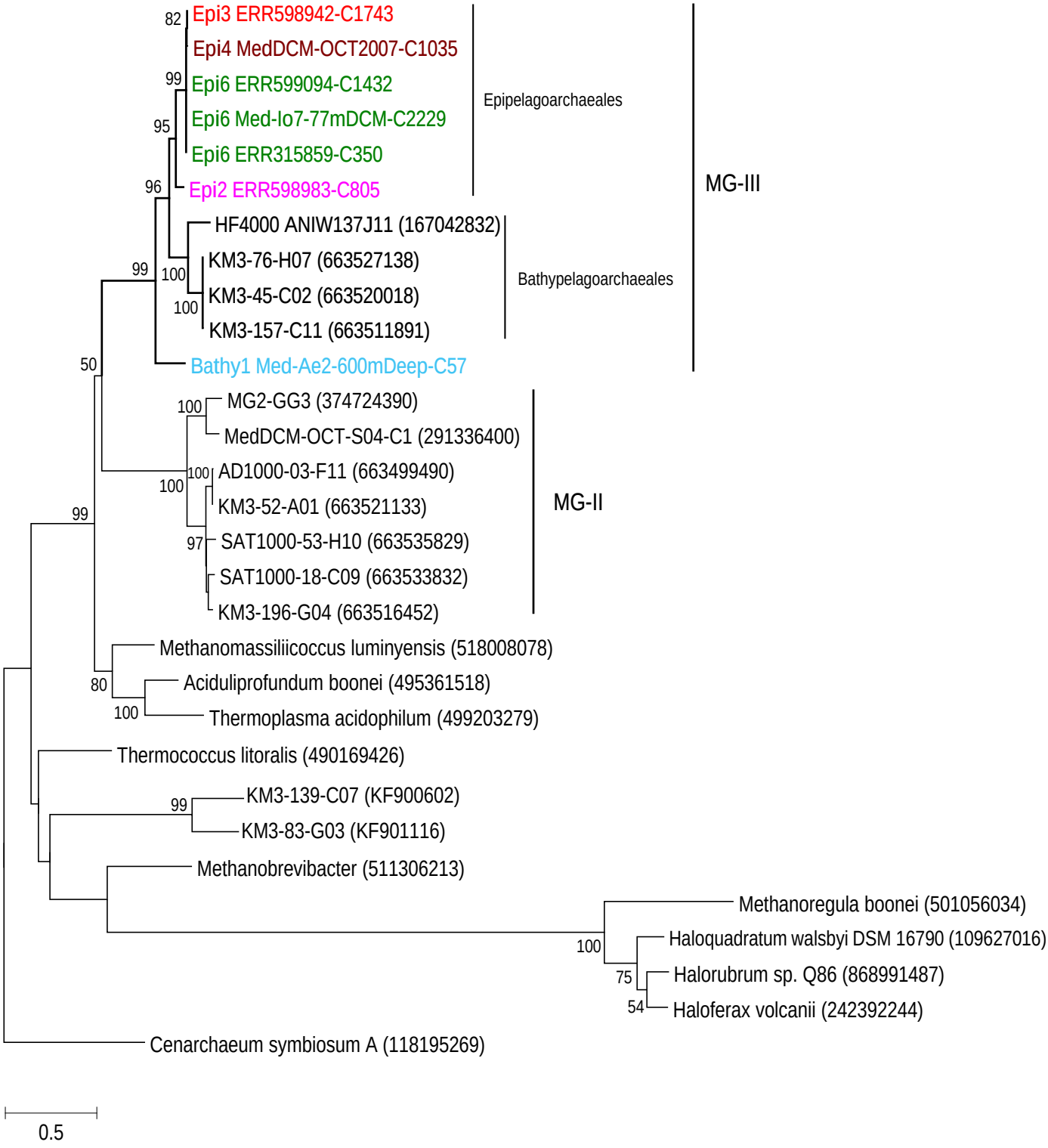
Supplementary Figure S5



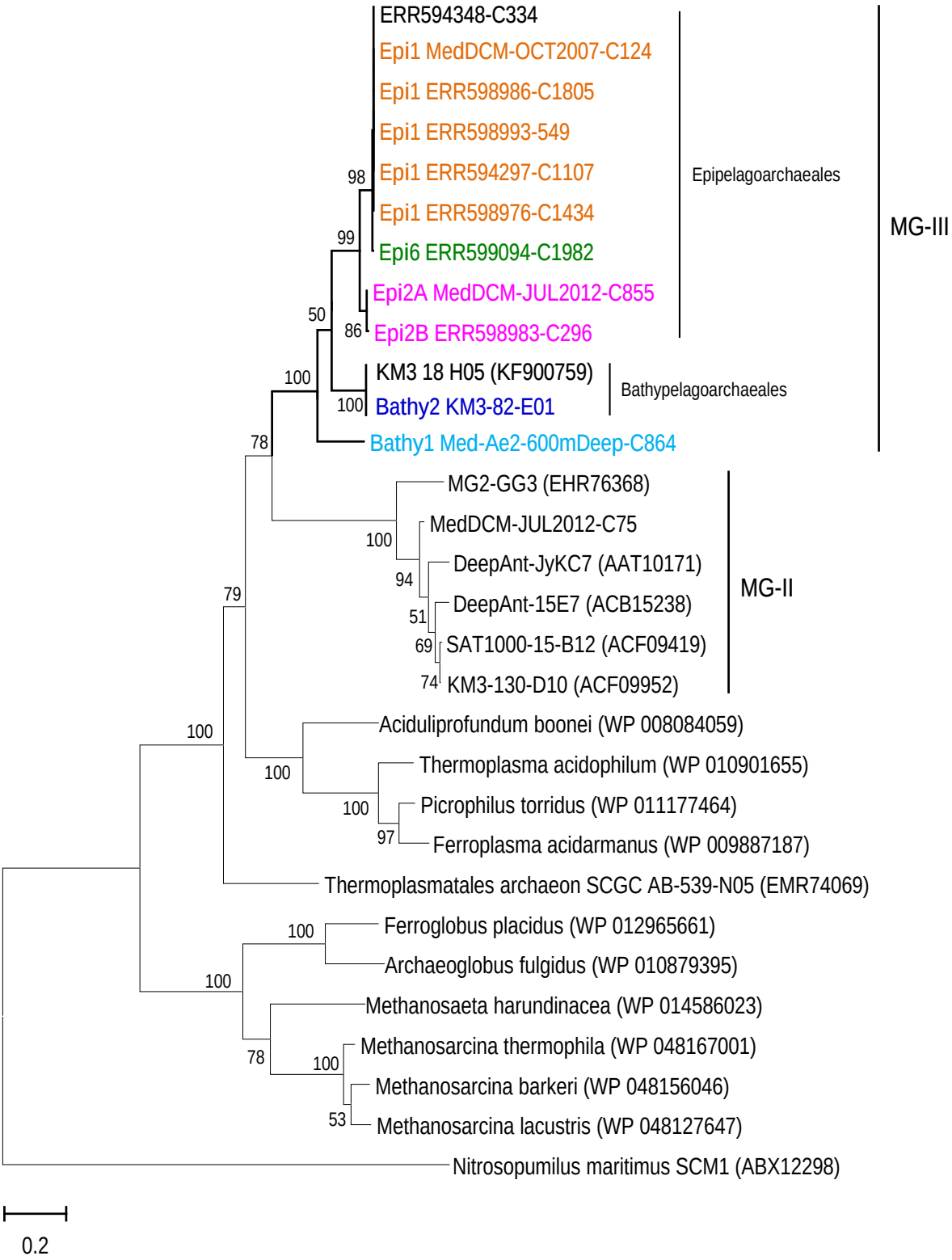
Supplementary Figure S6



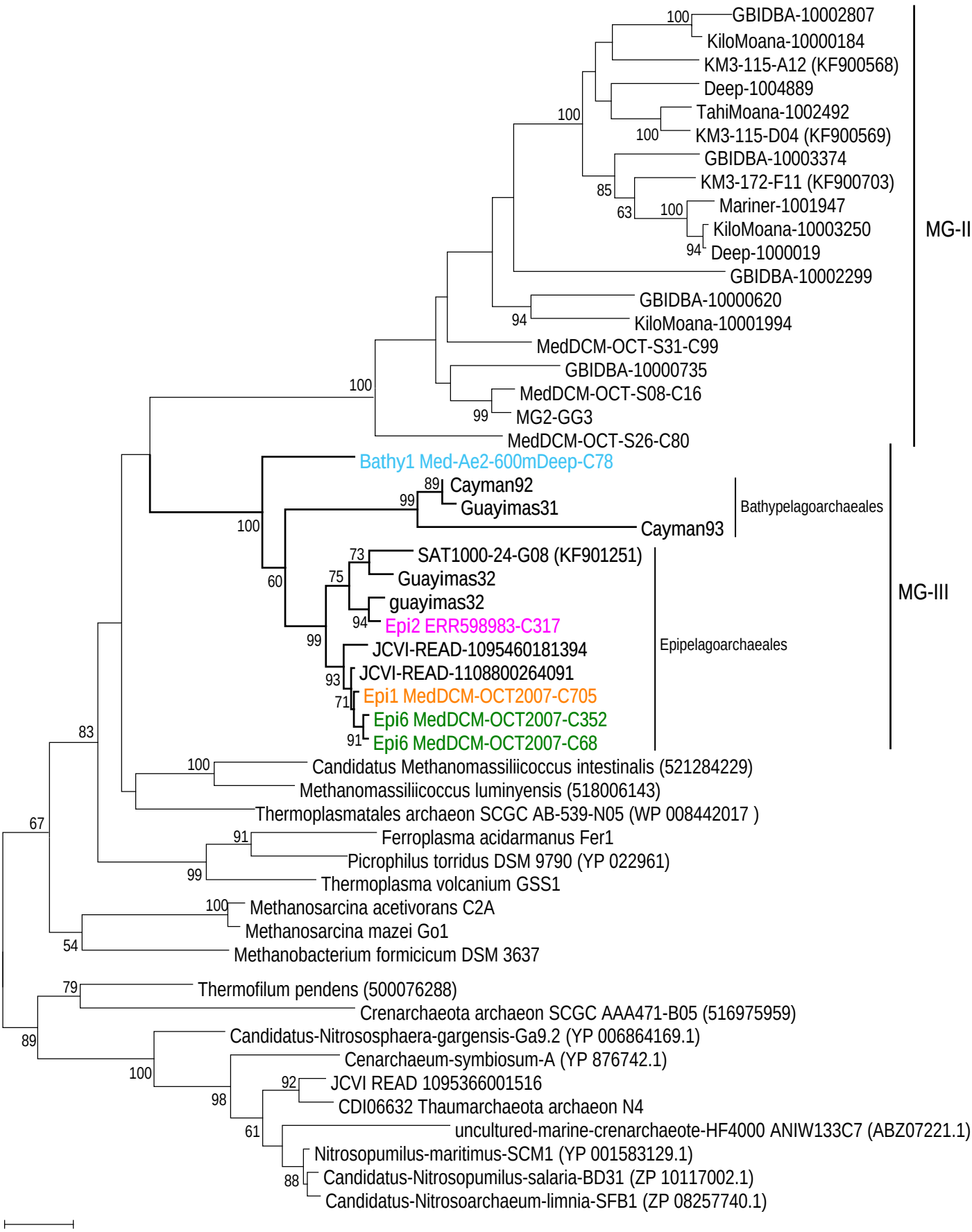
Supplementary Figure S7



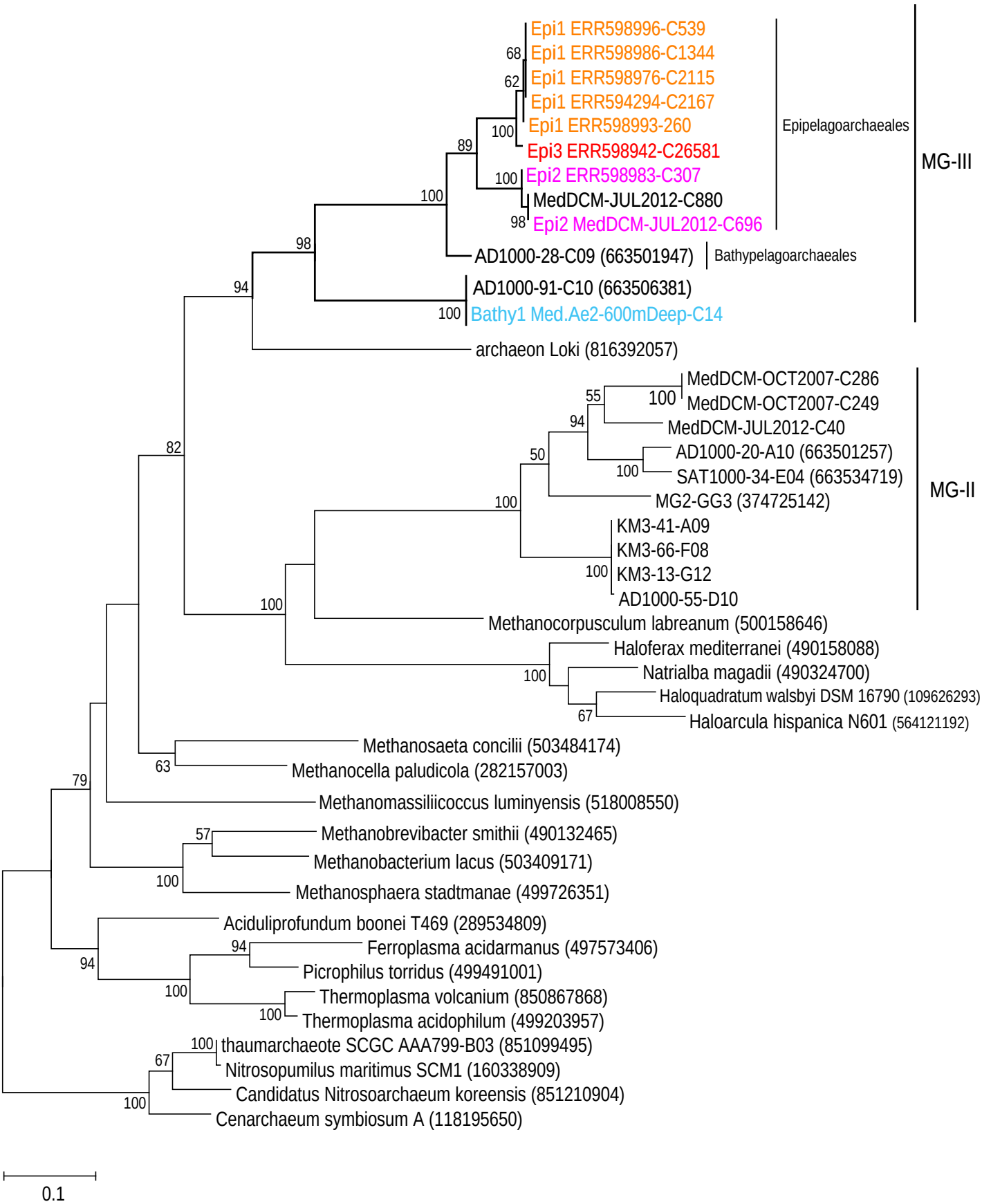
Supplementary Figure S8



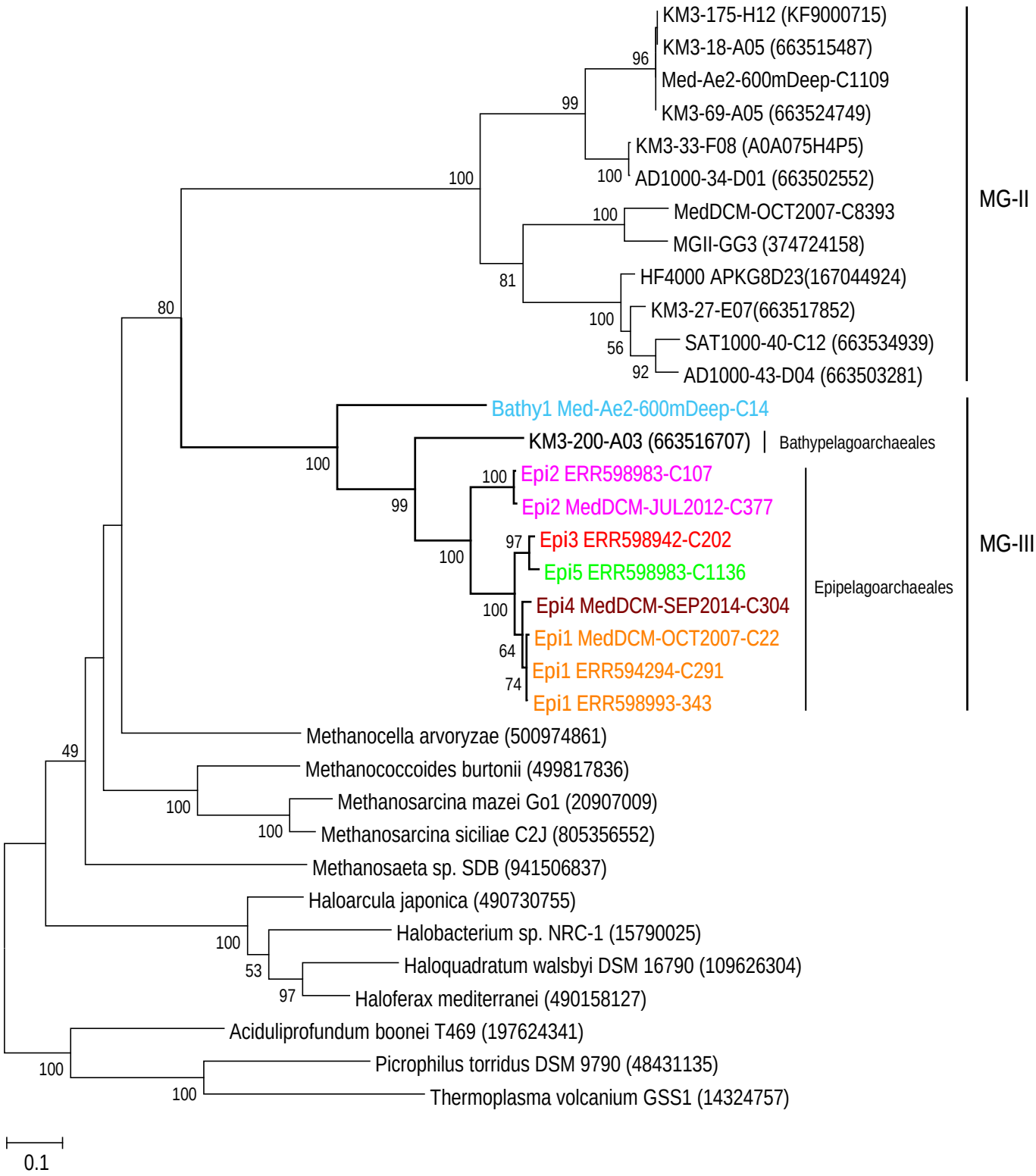
Supplementary Figure S9



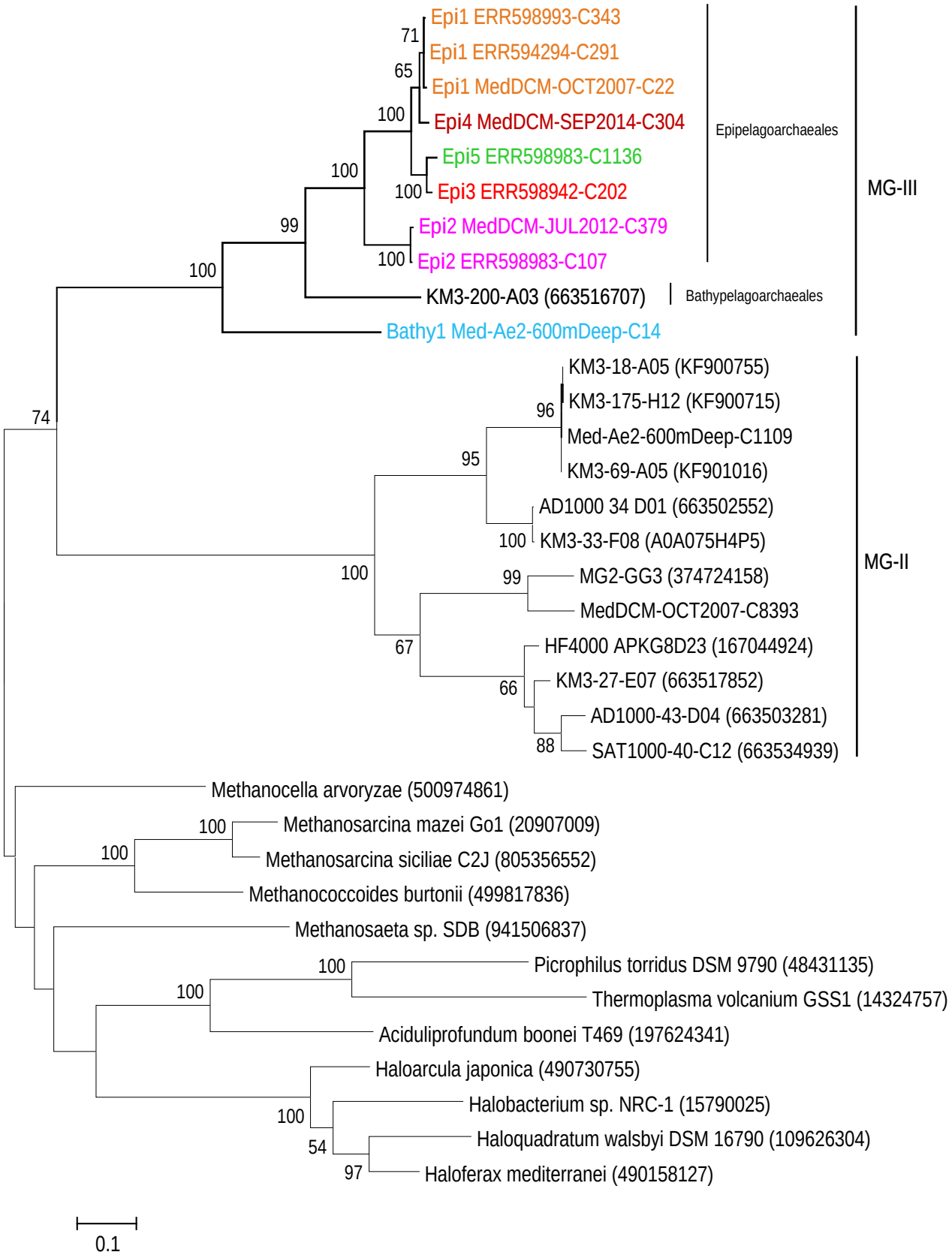
Supplementary Figure S10



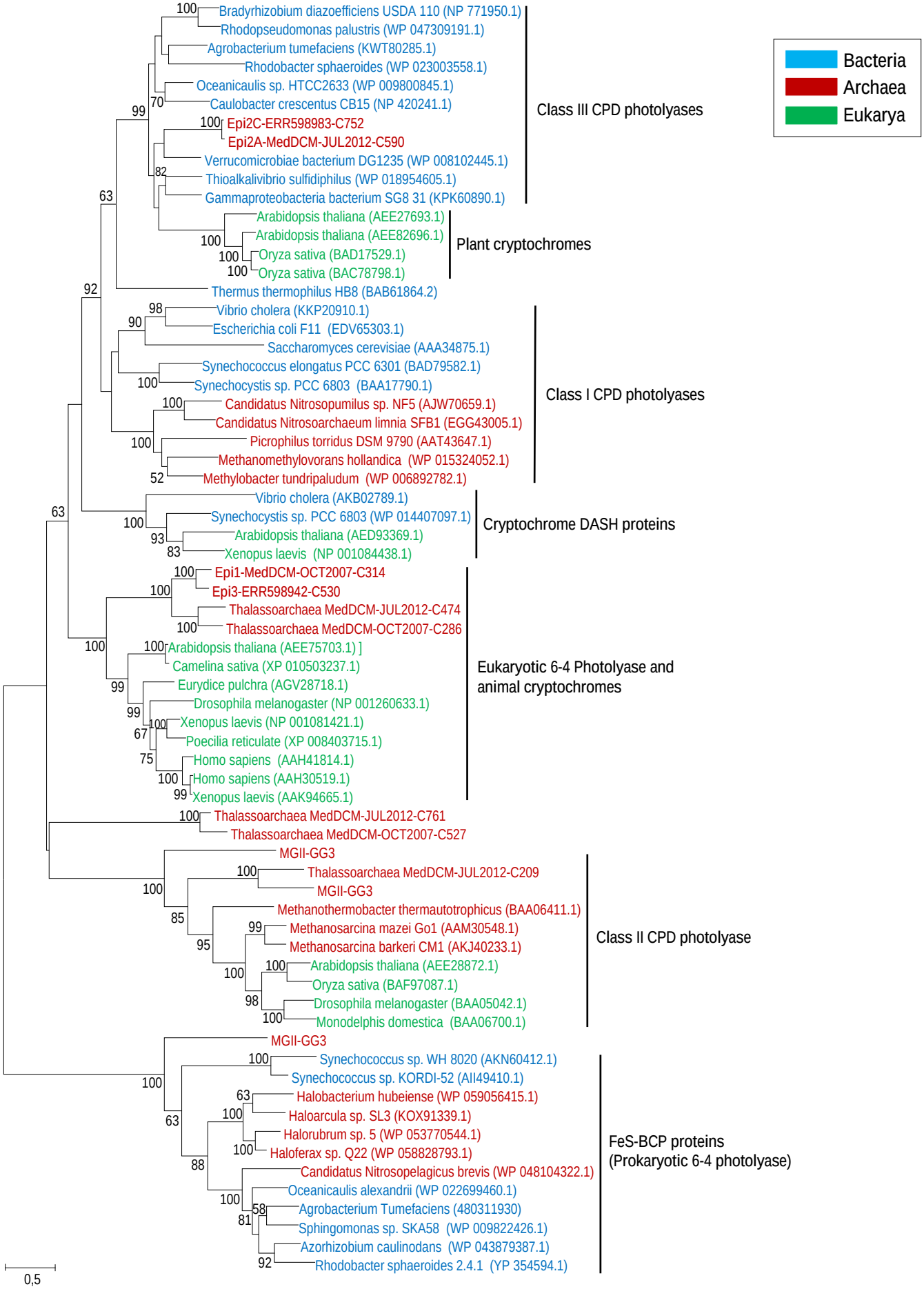
Supplementary Figure S11



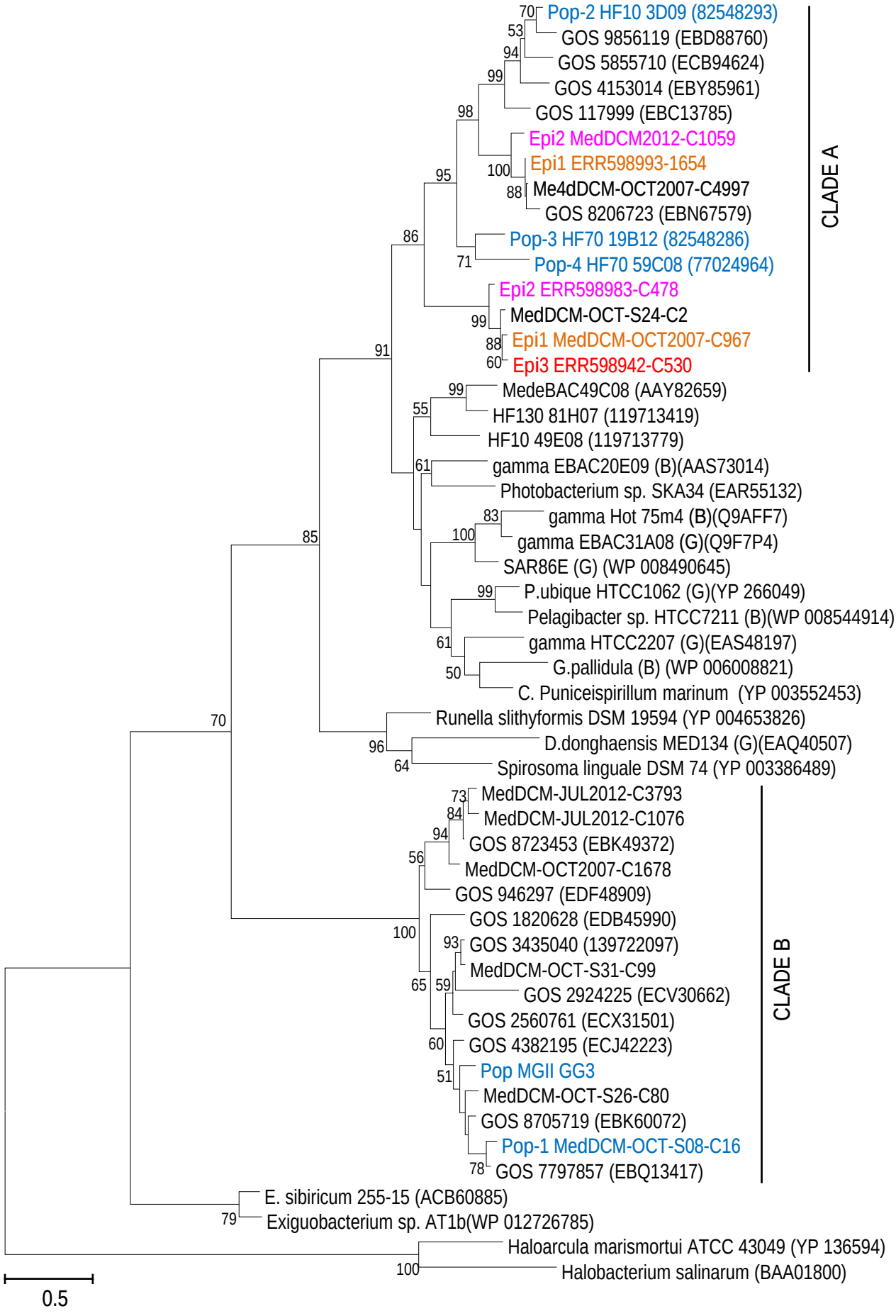
Supplementary Figure S12



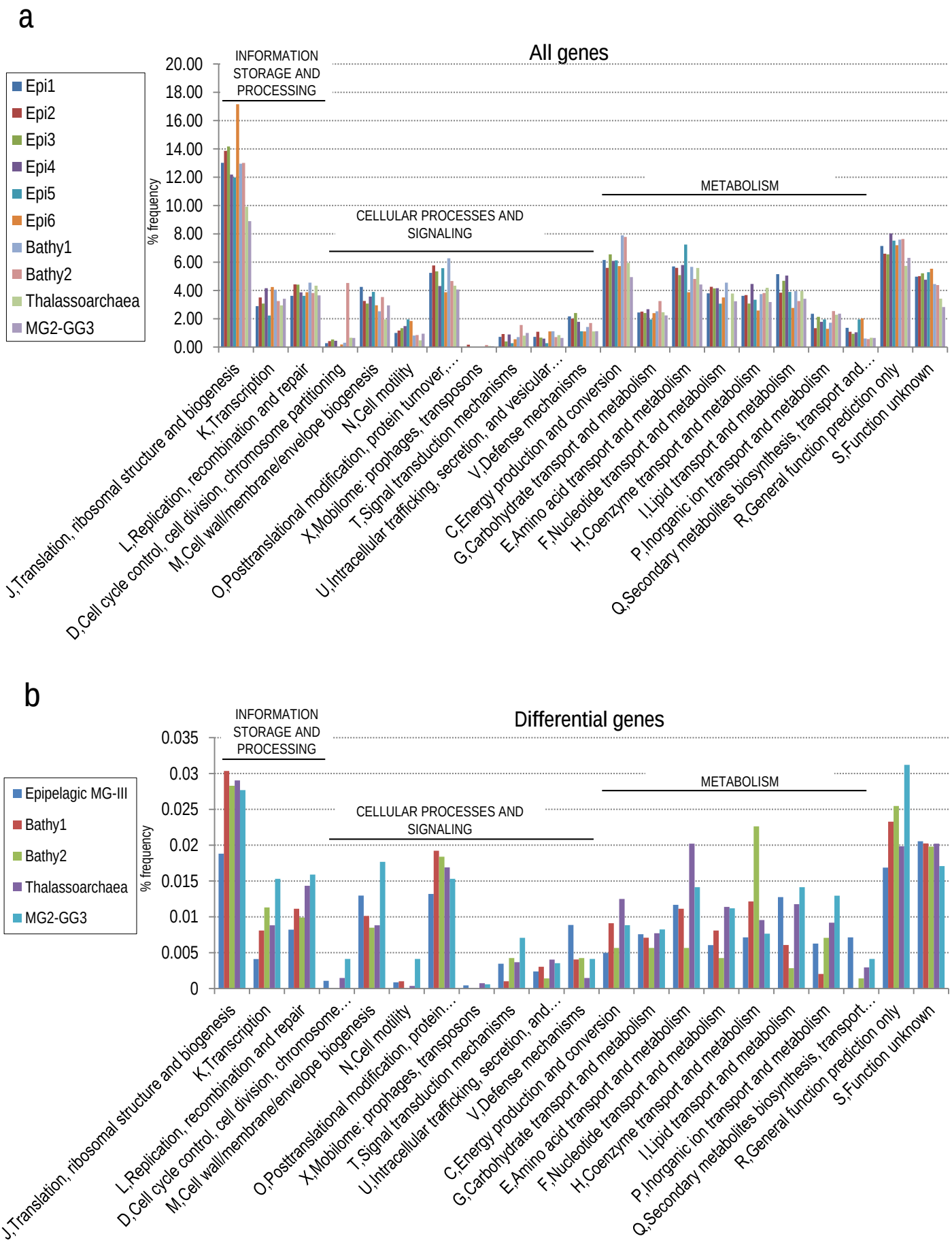
Supplementary Figure S13



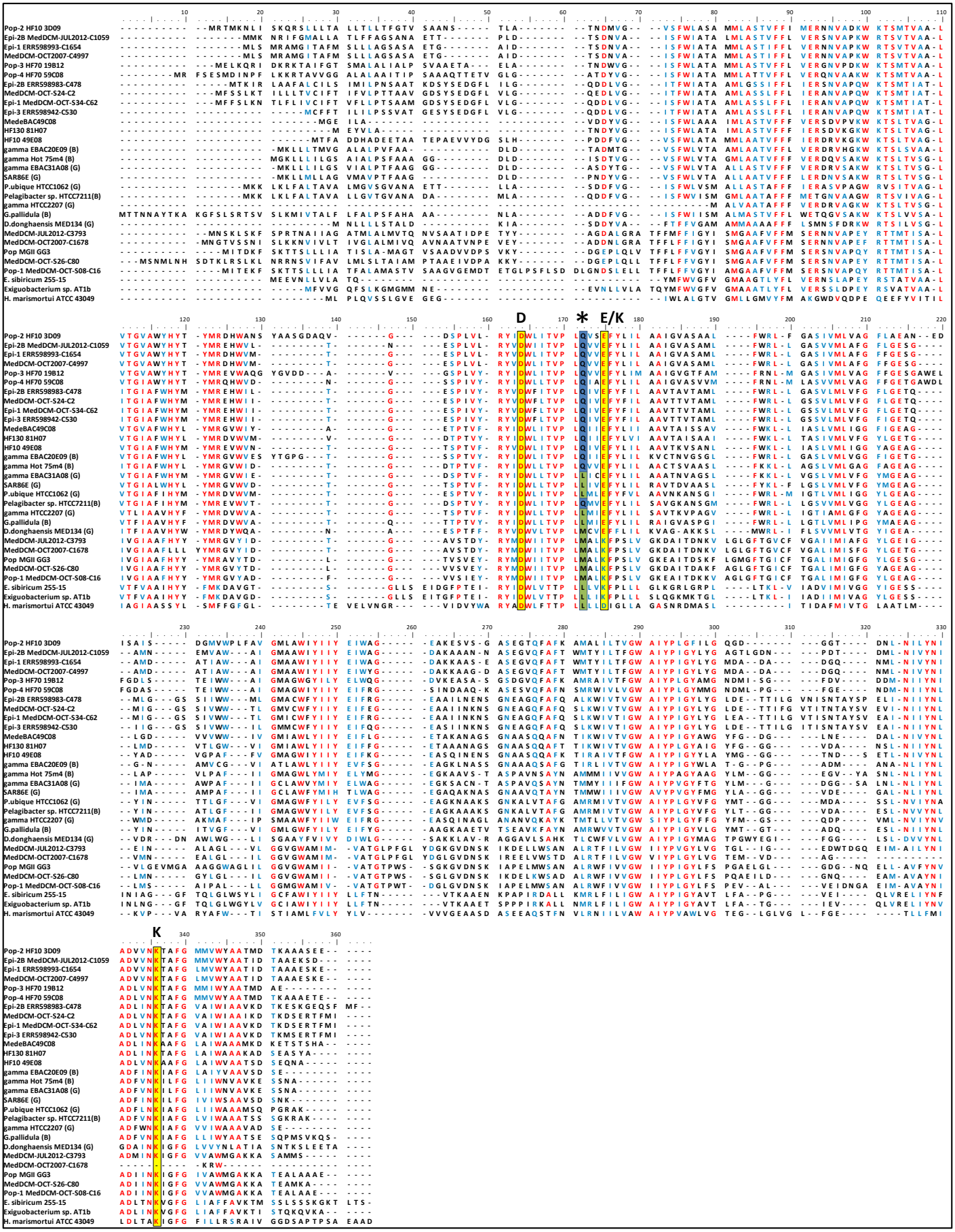
Supplementary Figure S14



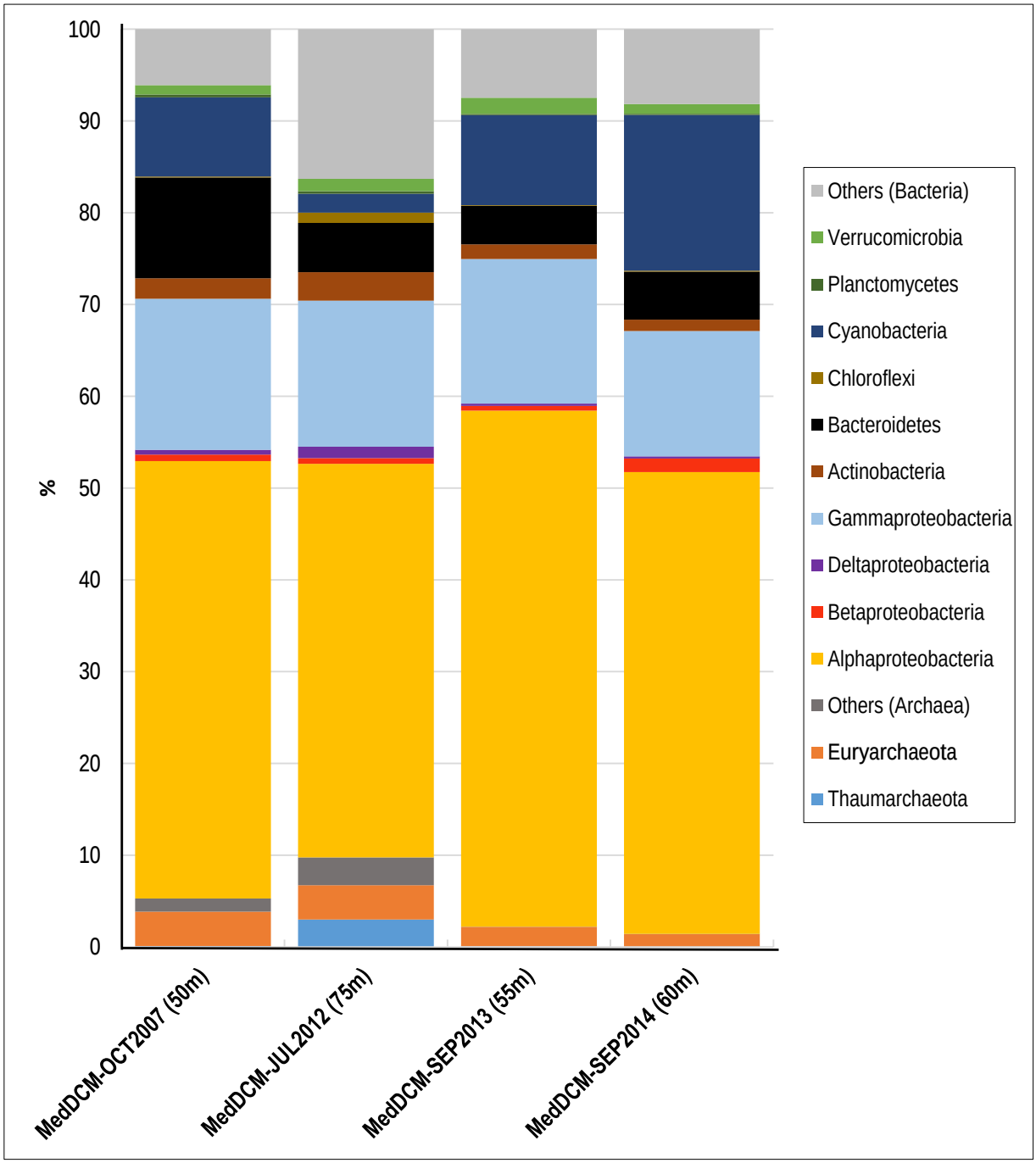
Supplementary Figure S15



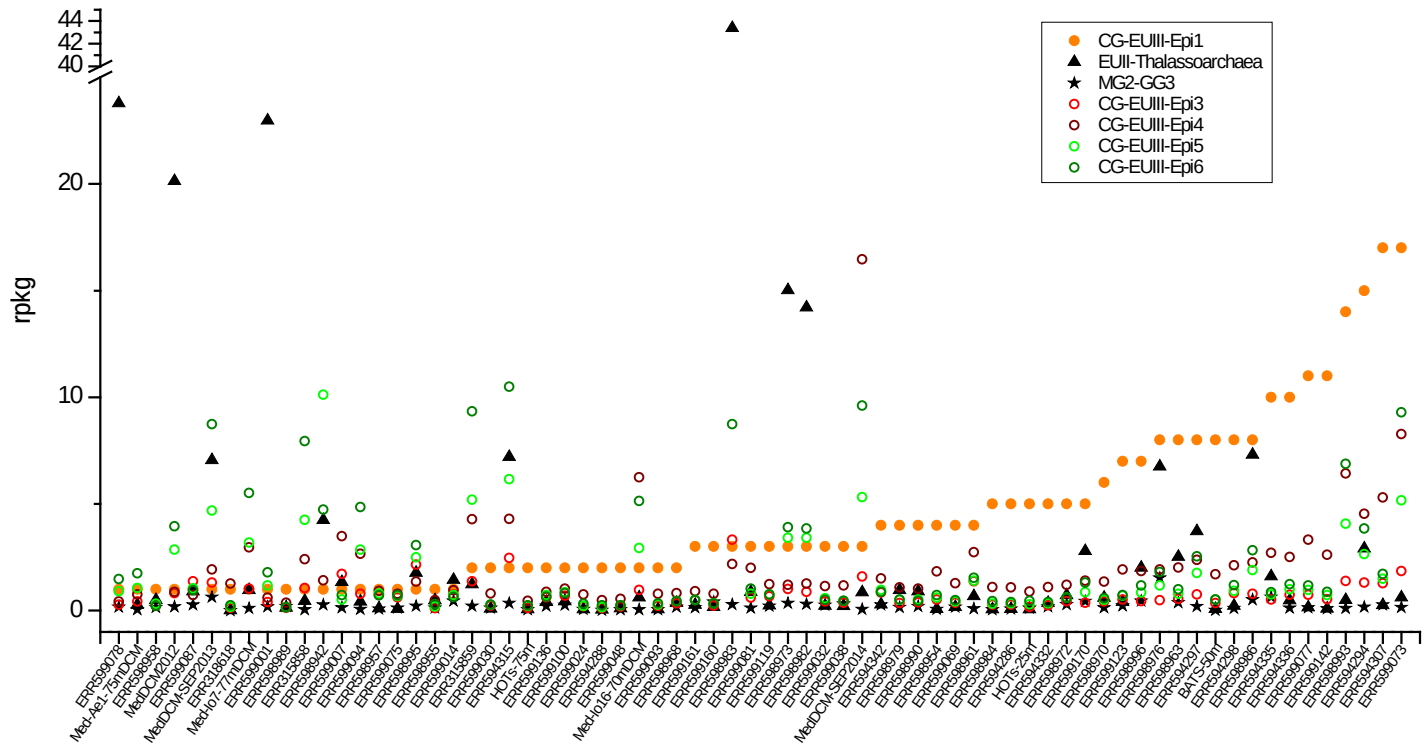
Supplementary Figure S16



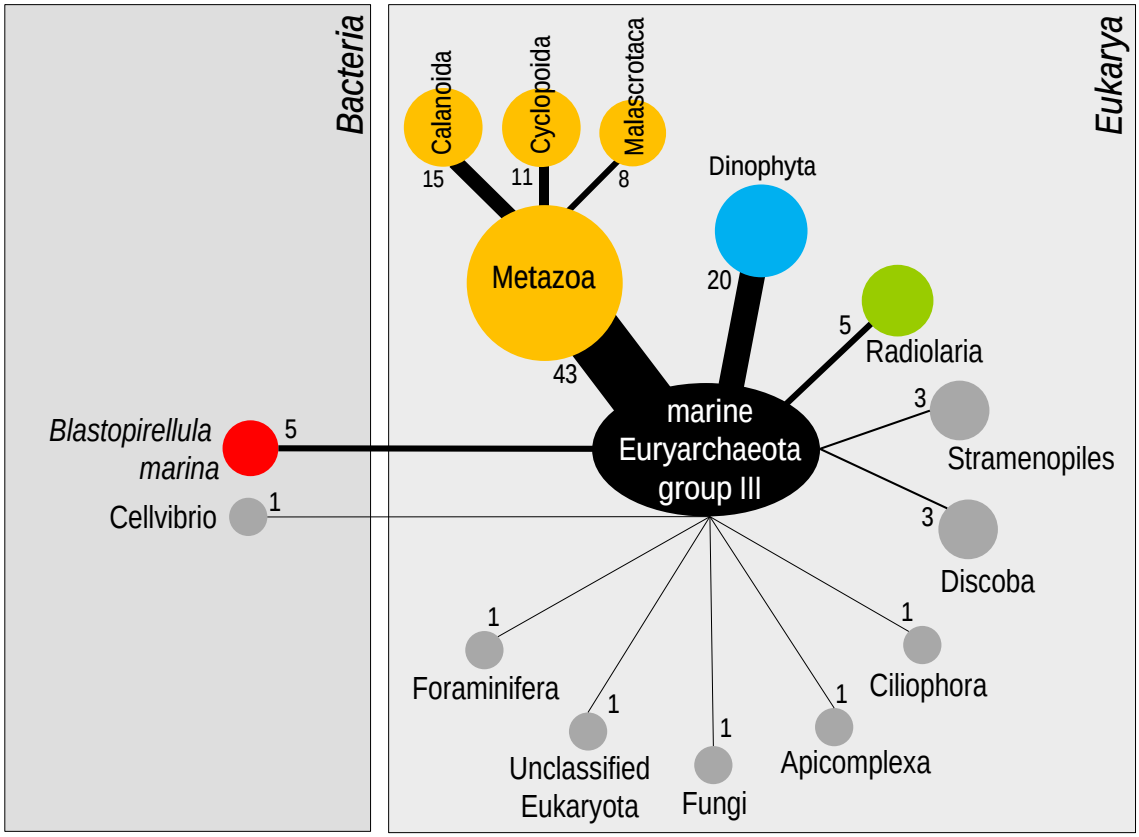
Supplementary Figure S17



Supplementary Figure S18



Supplementary Figure S19



Supplementary Table 1. List of contigs of each of the MG-III sequence bins

EPI-1					
Contig	GC	length (bp)	Contig	GC	length (bp)
ERR594294-C258	37.16	37101	ERR598986-C1752	36.81	10318
ERR594294-C291	35.99	33760	ERR598986-C1805	37.82	10158
ERR594294-C495	36.13	24149	ERR598993-C136	36.64	54729
ERR594294-C597	37.63	22103	ERR598993-C185	36.8	46411
ERR594294-C618	36.61	21634	ERR598993-C260	36.68	38768
ERR594294-C764	35.9	19337	ERR598993-C340	36.02	32610
ERR594294-C928	37.01	17143	ERR598993-C343	35.57	32391
ERR594294-C1053	36.67	15965	ERR598993-C353	36.78	31966
ERR594294-C1074	35.19	15847	ERR598993-C391	36.62	29994
ERR594294-C1114	37.11	15383	ERR598993-C404	36.3	29492
ERR594294-C1164	37.48	15019	ERR598993-C549	37.8	25250
ERR594294-C1249	36.35	14354	ERR598993-C566	36.12	24705
ERR594294-C1343	36.5	13728	ERR598993-C612	36.33	23521
ERR594294-C1553	38.3	12576	ERR598993-C652	36.6	22734
ERR594294-C1639	36.45	12240	ERR598993-C814	36.15	20075
ERR594294-C1668	36.28	12116	ERR598993-C863	36.95	19611
ERR594294-C1732	37.13	11778	ERR598993-C1010	35.74	18057
ERR594294-C1944	37.42	10969	ERR598993-C1033	35.33	17827
ERR594294-C1971	35.58	10907	ERR598993-C1094	36	17255
ERR594294-C2015	36.31	10770	ERR598993-C1183	36.43	16497
ERR594294-C2167	37.1	10336	ERR598993-C1197	35.45	16346
ERR594294-C2278	36.01	10029	ERR598993-C1499	35.65	14401
ERR594297-C1107	38.74	14137	ERR598993-C1654	35.49	13613
ERR594297-C1723	37.13	11194	ERR598993-C1784	37.25	13048
ERR594335-C415	36.69	28852	ERR598993-C1876	35.41	12706
ERR594335-C1656	37.67	13091	ERR598993-C1988	37.09	12339
ERR594348-C334	37.71	21238	ERR598993-C2355	37.35	11142
ERR594348-C443	36.93	18426	ERR598993-C2381	36.42	11076
ERR594348-C623	37.34	15553	ERR598993-C2473	37.21	10870
ERR594348-C739	35.29	14481	ERR598993-C2653	35.57	10434
ERR594348-C758	36.8	14338	ERR598993-C2753	35.96	10196
ERR594348-C837	34.51	13773	ERR598993-C2785	37.11	10145
ERR594348-C893	35.26	13263	ERR598993-C2847	36.96	10045
ERR594348-C1015	35.29	12441	ERR598996-C297	35.91	23347
ERR594348-C1040	37.03	12263	ERR598996-C465	36.17	18389
ERR594348-C1055	37.53	12166	ERR598996-C539	37.01	17111
ERR594348-C1205	38.2	11346	ERR598996-C570	37.4	16664
ERR598976-C379	36.29	24549	ERR598996-C611	35.25	16045
ERR598976-C428	35.93	23336	ERR598996-C890	36.06	13204
ERR598976-C538	35.63	21149	ERR598996-C966	35.93	12697
ERR598976-C738	35.61	18125	ERR598996-C1413	37.06	10508
ERR598976-C803	36.28	17350	ERR599073-C63	35.62	16990
ERR598976-C831	37.18	17084	ERR599073-C79	35.64	15397
ERR598976-C834	37.48	17065	ERR599073-C169	33.59	11510
ERR598976-C1434	38.58	12960	ERR599073-C205	37.35	10667
ERR598976-C1543	37.32	12491	MedDCM-OCT2007-C22	34.38	120849
ERR598976-C1550	34.58	12463	MedDCM-OCT2007-C13	35.58	104363
ERR598976-C1600	35.26	12288	MedDCM-OCT2007-C38	35.67	98497
ERR598976-C1618	37.01	12210	MedDCM-OCT2007-C79	36.91	88977
ERR598976-C1798	37.3	11559	MedDCM-OCT2007-C841	35.31	52149
ERR598976-C2115	36.74	10626	MedDCM-OCT2007-C112	35.94	44426
ERR598976-C2157	37.65	10530	MedDCM-OCT2007-C448	36.93	41300
ERR598976-C2167	36.22	10510	MedDCM-OCT2007-C699	37.64	38571
ERR598976-C2315	36.84	10192	MedDCM-OCT2007-C325	36.67	38037
ERR598976-C2324	36.04	10168	MedDCM-OCT2007-C124	37.34	37963
ERR598986-C308	37.44	24737	MedDCM-OCT2007-C898	36.42	37849
ERR598986-C357	35.74	23350	MedDCM-OCT2007-C100	36.79	36482
ERR598986-C457	35.74	20499	MedDCM-OCT2007-C4076	37.04	34001
ERR598986-C703	35.9	16494	MedDCM-OCT2007-C511	35.2	33568
ERR598986-C912	35.12	14518	MedDCM-OCT2007-C967	36.91	33097
ERR598986-C980	34.91	13981	MedDCM-OCT2007-C314	37.48	32394
ERR598986-C1168	37.57	12752	MedDCM-OCT2007-C222	36.42	32234
ERR598986-C1198	37.1	12565	MedDCM-OCT2007-C334	38.45	28213
ERR598986-C1274	37.34	12163	MedDCM-OCT2007-C1530	36.59	27625
ERR598986-C1293	35.29	12066	MedDCM-OCT2007-C705	37.38	22218
ERR598986-C1344	37.48	11801	MedDCM-OCT2007-C306	37.41	20163
ERR598986-C1671	36.33	10564	MedDCM-OCT2007-C818	38.84	16170
ERR598986-C1685	37.79	10503	MedDCM-OCT2007-C676	37.17	15511

EPI-2A		
Contig	GC	length (bp)
MedDCM-JUL2012-C377	35.98	27364
MedDCM-JUL2012-C368	34.94	24365
MedDCM-JUL2012-C409	36.5	22638
MedDCM-JUL2012-C516	36.57	22616
MedDCM-JUL2012-C803	36.14	22030
MedDCM-JUL2012-C602	35.56	21468
MedDCM-JUL2012-C1228	36.1	21053
MedDCM-JUL2012-C619	36.26	20185
MedDCM-JUL2012-C643	36.8	19543
MedDCM-JUL2012-C588	35.27	17776
MedDCM-JUL2012-C590	36.18	17749
MedDCM-JUL2012-C747	35.43	17605
MedDCM-JUL2012-C675	36.46	16175
MedDCM-JUL2012-C686	36.58	16035
MedDCM-JUL2012-C696	35.75	15875
MedDCM-JUL2012-C1445	37.44	15686
MedDCM-JUL2012-C1428	36.31	15645
MedDCM-JUL2012-C768	34.89	14974
MedDCM-JUL2012-C983	35.79	14911
MedDCM-JUL2012-C833	34.95	14179
MedDCM-JUL2012-C840	37.19	14037
MedDCM-JUL2012-C855	38.3	13838
MedDCM-JUL2012-C863	36.84	13763
MedDCM-JUL2012-C898	34.81	13277
MedDCM-JUL2012-C1480	34.3	12843
MedDCM-JUL2012-C1342	35.98	12072
MedDCM-JUL2012-C1407	36.71	11597
MedDCM-JUL2012-C1485	33.85	11279
MedDCM-JUL2012-C1587	36.42	10707
MedDCM-JUL2012-C1294	36.93	10398
MedDCM-JUL2012-C1295	35.14	10394
MedDCM-JUL2012-C1698	35.26	10393
MedDCM-JUL2012-C1741	34.99	10196
MedDCM-JUL2012-C1353	34.91	10030

EPI-2B		
Contig	GC	length (bp)
ERR598983-C49	36.19	79058
ERR598983-C129	36.05	53682
ERR598983-C296	37.62	33771
ERR598983-C317	36.37	32843
ERR598983-C403	36.72	28762
ERR598983-C471	36.1	26223
ERR598983-C478	37.47	26021
ERR598983-C805	36.56	20045
ERR598983-C942	36.08	18382
ERR598983-C1021	36.56	17585
ERR598983-C1043	37.58	17390
ERR598983-C1212	36.52	16024
ERR598983-C1565	34.57	14057
ERR598983-C1805	36.9	12954
ERR598983-C1824	37.45	12905
ERR598983-C1894	34.64	12704
MedDCM-JUL2012-C1134	36.6	23874
MedDCM-JUL2012-C1098	37.47	18009
MedDCM-JUL2012-C1192	35.87	16926
MedDCM-JUL2012-C1100	35.28	12974
MedDCM-JUL2012-C1185	35.92	12194
MedDCM-JUL2012-C1029	35.85	12049
MedDCM-JUL2012-C1059	34.73	11824
MedDCM-JUL2012-C1067	35.47	11769
MedDCM-JUL2012-C1126	36.15	11328
MedDCM-JUL2012-C1128	33.9	11309

EPI-2C		
Contig	GC	length (bp)
ERR598983-C29	35.29	92834
ERR598983-C264	36.35	36175
ERR598983-C307	37.33	33440
ERR598983-C415	36.03	28285
ERR598983-C455	33.96	26756
ERR598983-C752	36.4	20887
ERR598983-C870	36.1	19075
ERR598983-C1176	36.55	16240
ERR598983-C1232	35	15953
ERR598983-C1333	35.56	15364

EPI-3					
Contig	GC	length (bp)	Contig	GC	length (bp)
ERR598942-C113	35.54	55782	ERR598942-C1076	36.96	17164
ERR598942-C202	34.84	43429	ERR598942-C1159	36.43	16478
ERR598942-C302	35.53	35446	ERR598942-C1332	35.61	15132
ERR598942-C339	35.82	33474	ERR598942-C1391	35.42	14753
ERR598942-C445	34.65	28224	ERR598942-C1396	34.98	14733
ERR598942-C449	35.23	27959	ERR598942-C1476	36.49	14312
ERR598942-C481	36.93	27084	ERR598942-C1503	37.49	14168
ERR598942-C494	35.35	26753	ERR598942-C1625	36.29	13546
ERR598942-C530	35.95	25582	ERR598942-C1641	36.61	13491
ERR598942-C531	35.88	25568	ERR598942-C1743	37.19	13056
ERR598942-C563	34.34	24908	ERR598942-C1776	38.68	12920
ERR598942-C588	36.71	24366	ERR598942-C1797	35.55	12827
ERR598942-C669	36.9	22564	ERR598942-C2010	35.3	11945
ERR598942-C836	35.43	19885	ERR598942-C2020	35.47	11916
ERR598942-C852	33.95	19671	ERR598942-C2268	38.12	11189
ERR598942-C899	36.04	19191	ERR598942-C2464	35.03	10696
ERR598942-C963	36.41	18534	ERR598942-C2658	36.25	10233

EPI-4		
Contig	GC	length (bp)
MedDCM-SEP2014-C87	37.04	47235
MedDCM-SEP2014-C128	36.26	40668
MedDCM-SEP2014-C148	37	39153
MedDCM-SEP2014-C247	36.71	31052
MedDCM-SEP2014-C304	36.44	27517
MedDCM-SEP2014-C307	36.43	27494
MedDCM-SEP2014-C383	34.72	24850
MedDCM-SEP2014-C411	35.59	24120
MedDCM-SEP2014-C552	35.47	20187
MedDCM-SEP2014-C554	37.18	20122
MedDCM-SEP2014-C667	37.41	18345
MedDCM-SEP2014-C672	36.21	18315
MedDCM-SEP2014-C690	37.48	18089
MedDCM-SEP2014-C703	37.68	17936
MedDCM-SEP2014-C793	36.1	17073
MedDCM-SEP2014-C803	36.77	16955
MedDCM-SEP2014-C851	36.55	16353
MedDCM-SEP2014-C857	34.82	16283
MedDCM-SEP2014-C888	36.76	16025
MedDCM-SEP2014-C940	37.03	15611
MedDCM-SEP2014-C1047	36.5	14538
MedDCM-SEP2014-C1069	36.87	14335
MedDCM-SEP2014-C1118	35.95	13922
MedDCM-SEP2014-C1286	35.57	12756
MedDCM-SEP2014-C1299	37.61	12697
MedDCM-SEP2014-C1327	35.98	12543
MedDCM-SEP2014-C1485	37.54	11717
MedDCM-SEP2014-C1762	35.48	10580
MedDCM-SEP2014-C1880	36.01	10193
MedDCM-SEP2014-C1925	34.68	10057
MedDCM-OCT2007-C1035	36.78	34544

EPI-5		
Contig	GC	length (bp)
ERR598942-C1146	35.6	16544
ERR598942-C2164	35.9	11461
ERR598942-C2190	35.24	11409
ERR598942-C2311	35.88	11050
ERR598942-C2561	38.59	10464
ERR598983-C594	35.52	23368
ERR598983-C718	36.85	21372
ERR598983-C975	36.27	17947
ERR598983-C1136	35.76	16520
ERR598983-C1174	35.36	16242
ERR598983-C1305	36.21	15490
ERR598983-C1450	36.43	14690
ERR598983-C1599	36.4	13848
ERR598983-C1619	37.02	13764
ERR598983-C1680	35.91	13425
ERR598983-C2504	36.19	10823
ERR598983-C2521	35.37	10793
ERR598983-C2624	35.3	10525

EPI-6		
Contig	GC	length (bp)
ERR315859-C350	37.38	11676
ERR599094-C1432	37.42	13606
ERR599094-C1899	34.8	11443
ERR599094-C1956	35.97	11228
ERR599094-C1982	37.15	11155
MedDCM-OCT2007-C68	36.51	111071
MedDCM-OCT2007-C633	34.28	55596
MedDCM-OCT2007-C188	36.93	47138
MedDCM-OCT2007-C22970	35.34	42173
MedDCM-OCT2007-C377	36.78	41956
MedDCM-OCT2007-C3050	35.91	41814
MedDCM-OCT2007-C824	36.05	41493
MedDCM-OCT2007-C352	36.98	40323
MedDCM-OCT2007-C449	37.06	39898
MedDCM-OCT2007-C1482	36.24	39690
MedDCM-OCT2007-C359	37.25	37219
MedDCM-OCT2007-C359	37.25	37219
MedDCM-OCT2007-C1276	36.23	36777
MedDCM-OCT2007-C368	37.37	34684
MedDCM-OCT2007-C360	36.86	33348
MedDCM-OCT2007-C174	36.8	32379
MedDCM-OCT2007-C2146	36.56	21038
Med-Io7-77mDCM-C1113	36.16	18506
Med-Io7-77mDCM-C1303	35.46	13394
Med-Io7-77mDCM-C2049	35.01	10014
Med-Io7-77mDCM-C2229	36.64	13869

BATHY-1		
Contig	GC	length (bp)
AD1000-41-D11	35.88	36746
AD1000-61-A07	36.68	30250
AD1000-91-C10	35.96	38736
KM3-03-G06-C1	36.91	34695
KM3-03-G06-C2	35.77	33289
KM3-12-H02-C01	38.92	36194
KM3-12-H02-C02	36.56	35952
KM3-147-B09	36.94	37598
Med-Ae2-600mDeep-C14	36.85	94706
Med-Ae2-600mDeep-C40	38.62	67903
Med-Ae2-600mDeep-C47	36.05	63835
Med-Ae2-600mDeep-C50	35.64	63144
Med-Ae2-600mDeep-C53	37.02	61855
Med-Ae2-600mDeep-C57	37.81	59790
Med-Ae2-600mDeep-C62	36.05	52612
Med-Ae2-600mDeep-C78	36.86	47080
Med-Ae2-600mDeep-C83	36.31	45325
Med-Ae2-600mDeep-C97	38.33	41337
Med-Ae2-600mDeep-C181	38.56	28692
Med-Ae2-600mDeep-C194	37.74	28187
Med-Ae2-600mDeep-C277	36.5	23394
Med-Ae2-600mDeep-C336	36.75	20947
Med-Ae2-600mDeep-C344	36.81	20732
Med-Ae2-600mDeep-C488	36.78	17189
Med-Ae2-600mDeep-C599	37.9	15102
Med-Ae2-600mDeep-C647	35.65	14260
Med-Ae2-600mDeep-C738	38.23	13111
Med-Ae2-600mDeep-C748	37.3	13028
Med-Ae2-600mDeep-C864	37.72	11932
Med-Ae2-600mDeep-C875	37.08	11813
Med-Ae2-600mDeep-C898	36.79	11640
Med-Ae2-600mDeep-C921	36.66	11445
Med-Ae2-600mDeep-C954	36.12	11128
Med-Ae2-600mDeep-C966	37.49	11055
Med-Ae2-600mDeep-C999	38.76	10785
Med-Ae2-600mDeep-C1034	38.51	10528
Med-Ae2-600mDeep-C1094	38.3	10194
Med-Ae2-600mDeep-C1105	36.06	10167
Med-Ae2-600mDeep-C1120	37.54	10120

BATHY-2		
Contig	GC	length (bp)
KM3-110-C01	63.75	36614
KM3-141-A08	64.58	33669
KM3-148-H03	65.28	36104
KM3-152-H07	64.68	36691
KM3-156-D02	65.63	32803
KM3-157-C11	64.42	36794
KM3-161-F10	67.44	41692
KM3-164-G07	66.03	21529
KM3-176-D09	62.39	2430
KM3-177-A07	60.44	30978
KM3-177-C07	65.95	7771
KM3-18-H05	64.03	36289
KM3-181-H05	64.58	32150
KM3-185-B03	62.49	39315
KM3-188-A01	64.77	32963
KM3-190-A12	64.89	34623
KM3-191-F05	61.01	32444
KM3-200-A03	61.24	7346
KM3-202-G05	63.24	36284
KM3-205-C10	65.19	10201
KM3-33-C12	64.56	14809
KM3-35-H09	62.95	33996
KM3-37-D11	64.49	41030
KM3-43-F08	67.29	30905
KM3-45-C02	65.18	34134
KM3-51-F05	64.27	37422
KM3-54-A09	64.98	36109
KM3-59-B11	66.03	30687
KM3-61-H04	62.66	35580
KM3-65-G10	67.77	16442
KM3-75-F08	63.04	36690
KM3-76-H07	64.42	34062
KM3-79-B02	66.8	8021
KM3-82-E01	65.13	37648
KM3-89-F04	64.24	18265
KM3-91-D09	61.22	35105
SAT1000-51-E12	67.2	1829

NO CLASSIFIED CONTIGS		
Contig	GC	length (bp)
ERR315859-C1033	41.17	7278
ERR594307-C1518	40.19	11691
ERR594348-C334	37.71	21238
ERR598942-C4580	40.12	7492
ERR598983-C1332	37.14	15365
ERR598983-C447	36.88	27099
ERR598986-C1198	37.1	12565
ERR598993-C1154	38.67	16689
ERR598996-C4574	43.27	5775
ERR598996-C570	37.4	16664
ERR599170-C2618	41.63	6853
MedDCM-JUL2012-C379		
MedDCM-JUL2012-C880		
MedDCM-JUL2012-C882		
MedDCM-OCT2007-C765		
MedDCM-OCT2007-C967		

Supplementary Table 2. List of metagenomes used for recruitments in Figure 4a and Supplementary Figure 2B, sorted by temperature and depth.

Metagenome	Size fraction	Depth	Temperature (ºC)	Location	Work	Epi-1	Epi-3	Epi-4	Epi-5	Epi-6	Epi-2A	Epi-2B	Epi-2C	Bathy-1	Bathy-2	Cayman9	Cayman9	Guaymas32	Guaymas31	SCGCAAA-228E19	Thalassoarchaea	MG2-GG3	*A. boneii*	*N. maritimus*
ERR599075	0.22 µm	5m	30	Indian Monsoon Gyres Province	1	1.25	0.62	0.78	0.68	0.73	0.14	0.16	0.09	0.08	0.06	0.05	0.01	0.11	0.04	0.12	0.08	0.12	0.01	0.01
ERR599119	0.22 µm	5m	26.8	South Pacific Subtropical Gyre Province, North and South	1	3.02	0.67	1.24	0.76	0.75	0.16	0.19	0.11	0.10	0.10	0.09	0.01	0.14	0.09	0.29	0.24	0.21	0.03	0.01
ERR599030	0.22 µm	5m	26.6	North Pacific Equatorial Countercurrent Province	1	1.69	0.23	0.80	0.27	0.28	0.05	0.07	0.06	0.05	0.04	0.04	0.00	0.07	0.04	0.10	0.10	0.12	0.01	0.01
ERR599160	0.22 µm	5m	26.6	South Pacific Subtropical Gyre Province, North and South	1	2.66	0.23	0.79	0.34	0.34	0.06	0.06	0.06	0.05	0.04	0.04	0.00	0.06	0.05	0.21	0.17	0.41	0.03	0.01
ERR594307	0.22 µm	5m	26.5	South Pacific Subtropical Gyre Province, North and South	1	16.71	1.28	5.30	1.49	1.72	0.31	0.31	0.25	0.16	0.12	0.10	0.02	0.21	0.11	0.31	0.26	0.28	0.03	0.01
ERR599069	0.22 µm	5m	26.5	South Pacific Subtropical Gyre Province, North and South	1	4.31	0.36	1.28	0.43	0.48	0.09	0.09	0.07	0.06	0.05	0.05	0.00	0.08	0.05	0.16	0.17	0.17	0.02	0.02
ERR598989	0.22 µm	5m	26.4	North Pacific Equatorial Countercurrent Province	1	1.06	0.12	0.36	0.15	0.16	0.02	0.03	0.03	0.03	0.02	0.02	0.00	0.05	0.03	0.12	0.14	0.20	0.02	0.01
ERR599038	0.22 µm	5m	26.1	Pacific Equatorial Divergence Province	1	3.19	0.37	1.18	0.46	0.46	0.09	0.11	0.11	0.08	0.07	0.07	0.00	0.11	0.07	0.24	0.21	0.32	0.03	0.05
ERR599142	0.22 µm	5m	25.2	North Pacific Subtropical and Polar Front Provinces	1	11.42	0.53	2.61	0.72	0.87	0.12	0.13	0.09	0.07	0.05	0.05	0.01	0.08	0.04	0.10	0.10	0.10	0.01	0.00
ERR599093	0.22 µm	5m	25.1	South Pacific Subtropical Gyre Province, North and South	1	2.39	0.22	0.78	0.27	0.33	0.06	0.07	0.05	0.03	0.02	0.02	0.00	0.06	0.02	0.06	0.07	0.05	0.01	0.01
ERR598984	0.22 µm	5m	25	South Atlantic Gyral Province	1	4.59	0.28	1.10	0.32	0.44	0.06	0.07	0.03	0.03	0.02	0.01	0.00	0.05	0.02	0.06	0.07	0.05	0.01	0.00
ERR599136	0.22 µm	5m	25	Caribbean Province	1	2.05	0.62	0.87	0.72	0.65	0.31	0.39	0.16	0.22	0.15	0.15	0.02	0.30	0.14	0.46	0.41	0.21	0.03	0.06
ERR598954	0.22 µm	5m	24.2	South Pacific Subtropical Gyre Province, North and South	1	3.90	0.52	1.84	0.60	0.73	0.15	0.21	0.10	0.09	0.07	0.06	0.01	0.13	0.05	0.13	0.08	0.06	0.01	0.01
ERR594288	0.22 µm	5m	23.9	Mediterranean Sea, Black Sea Province	1	2.15	0.11	0.48	0.16	0.24	0.02	0.03	0.02	0.02	0.01	0.01	0.00	0.04	0.01	0.02	0.05	0.02	0.01	0.00
ERR599024	0.22 µm	5m	23.8	South Pacific Subtropical Gyre Province, North and South	1	2.08	0.20	0.75	0.22	0.32	0.05	0.06	0.03	0.03	0.02	0.02	0.00	0.05	0.03	0.06	0.06	0.07	0.01	0.01
ERR594286	0.22 µm	5m	23.3	South Atlantic Gyral Province	1	4.74	0.27	1.09	0.34	0.40	0.07	0.09	0.04	0.04	0.03	0.03	0.00	0.06	0.03	0.07	0.08	0.08	0.01	0.01
ERR599077	0.22 µm	5m	22.8	South Pacific Subtropical Gyre Province, North and South	1	11.16	0.74	3.32	0.96	1.17	0.17	0.19	0.15	0.09	0.08	0.06	0.00	0.12	0.07	0.16	0.16	0.14	0.01	0.00
ERR598970	0.22 µm	5m	22.2	Eastern Africa Coastal Province	1	5.77	0.42	1.35	0.50	0.57	0.10	0.12	0.06	0.06	0.05	0.04	0.01	0.08	0.05	0.15	0.60	0.14	0.01	0.03
ERR598979	0.22 µm	5m	21.8	Eastern Africa Coastal Province	1	3.83	0.35	1.08	0.50	0.52	0.09	0.10	0.07	0.07	0.06	0.06	0.01	0.11	0.06	0.23	0.97	0.15	0.02	0.07
ERR598993	0.22 µm	5m	21.4	Mediterranean Sea, Black Sea Province	1	13.70	1.39	6.42	4.07	6.87	1.02	0.76	0.45	0.10	0.07	0.06	0.01	0.17	0.07	0.19	0.50	0.11	0.01	0.10
ERR598955	0.22 µm	5m	20.5	North Atlantic Subtropical Gyral Province	1	1.42	0.11	0.49	0.18	0.31	0.01	0.02	0.03	0.02	0.01	0.01	0.00	0.04	0.01	0.05	0.49	0.13	0.01	0.01
ERR599123	0.22 µm	5m	20.4	North Atlantic Subtropical Gyral Province	1	6.62	0.51	1.92	0.69	0.72	0.25	0.27	0.12	0.13	0.09	0.07	0.01	0.20	0.08	0.26	0.41	0.23	0.03	0.34
ERR594332	0.22 µm	5m	19.9	South Atlantic Gyral Province	1	5.08	0.24	1.10	0.34	0.40	0.06	0.07	0.05	0.03	0.02	0.02	0.00	0.05	0.02	0.08	0.36	0.18	0.01	0.01
ERR594335	0.22 µm	5m	19.8	South Atlantic Gyral Province	1	9.71	0.51	2.70	0.76	0.86	0.14	0.17	0.08	0.11	0.07	0.06	0.01	0.13	0.07	0.35	1.61	0.63	0.04	0.02
ERR598968	0.22 µm	5m	19.1	North Atlantic Subtropical Gyral Province	1	2.44	0.34	0.81	0.44	0.44	0.32	0.33	0.14	0.16	0.08	0.07	0.01	0.23	0.07	0.24	0.37	0.17	0.03	0.79
ERR598963	0.22 µm	5m	18.7	North Atlantic Subtropical Gyral Province	1	7.76	0.54	2.01	0.78	1.00	0.35	0.29	0.15	0.13	0.09	0.09	0.01	0.17	0.08	0.34	2.52	0.38	0.02	0.46
ERR315858	0.22 µm	5m	17.6	Mediterranean Sea, Black Sea Province	1	1.21	1.06	2.40	4.25	7.94	0.43	0.36	0.25	0.05	0.04	0.02	0.00	0.13	0.04	0.05	0.46	0.05	0.01	0.06
ERR599170	0.22 µm	5m	17.6	North Atlantic Subtropical Gyral Province	1	5.25	0.37	1.40	0.86	1.31	0.20	0.18	0.11	0.05	0.03	0.03	0.00	0.09	0.03	0.16	2.79	0.44	0.02	0.10
ERR598976	0.22 µm	5m	17.3	North Atlantic Subtropical Gyral Province	1	7.70	0.49	1.92	1.17	1.80	0.15	0.15	0.08	0.06	0.05	0.05	0.00	0.08	0.06	0.24	6.76	1.54	0.03	0.02
ERR594297	0.22 µm	5m	16.8	South Atlantic Gyral Province	1	7.87	0.76	2.37	1.76	2.54	1.30	1.03	0.50	0.21	0.11	0.10	0.02	0.31	0.10	0.30	3.73	0.20	0.02	0.19
ERR598973	0.22 µm	5m	15	Benguela Current Coastal Province	1	3.05	1.01	1.21	3.41	3.91	1.01	0.81	0.67	0.07	0.03	0.03	0.00	0.11	0.03	0.15	15.02	0.35	0.01	0.51
ERR599078	0.22 µm	5m	14.3	North Atlantic Subtropical Gyral Province	1	0.55	0.28	0.43	0.90	1.47	0.78	0.47	0.35	0.04	0.03	0.03	0.01	0.09	0.03	0.18	23.79	0.17	0.01	0.29
ERR598983	0.22 µm	5m	14.1	Gulf Stream Province	1	2.72	3.32	2.18	32.10	8.74	5.25	9.11	9.81	0.17	0.13	0.11	0.01	0.49	0.10	0.35	43.40	0.29	0.03	2.39
ERR599032	0.22 µm	40m	26.1	Pacific Equatorial Divergence Province	1	3.18	0.42	1.15	0.57	0.50	0.13	0.14	0.08	0.08	0.07	0.05	0.01	0.12	0.07	0.25	0.23	0.26	0.03	0.06
HOTs	0.22 µm	25m	25	North Pacific Subtropical Gyre	5	5.05	0.17	0.89	0.29	0.45	0.09	0.17	0.12	0.04	0.00	0.00	0.00	0.09	0.03	0.03	0.06	0.12	0.01	0.00
BATS	0.22 µm	50m	25	North Atlantic	5	8.08	0.35	1.70	0.51	0.52	0.20	0.09	0.00	0.11	0.03	0.04	0.00	0.13	0.03	0.18	0.08	0.02	0.01	0.01
ERR598990	0.22 µm	30m	21.8	Eastern Africa Coastal Province	1	3.88	0.38	1.03	0.52	0.45	0.11	0.12	0.08	0.11	0.08	0.10	0.04	0.20	0.10	0.36	0.91	0.20	0.02	0.13
ERR599014	0.22 µm	50m	21.8	South Pacific Subtropical Gyre Province, North and South	1	1.45	0.76	0.95	0.72	0.64	0.12	0.16	0.10	0.08	0.07	0.07	0.01	0.11	0.07	0.29	1.44	0.46	0.03	0.02
ERR599081	0.22 µm	50m	20.6	South Pacific Subtropical Gyre Province, North and South	1	2.98	0.61	2.00	0.78	1.02	0.24	0.24	0.15	0.10	0.06	0.05	0.01	0.15	0.05	0.17	0.87	0.12	0.01	0.02
ERR599007	0.22 µm	40m	19.6	Pacific Equatorial Divergence Province	1	1.22	1.71	3.49	0.52	0.73	0.08	0.11	0.08	0.04	0.03	0.02	0.01	0.07	0.03	0.11	1.34	0.14	0.01	0.01
ERR598987	0.22 µm	40m	18.6	North Pacific Equatorial Countercurrent Province	1	0.11	0.13	0.07	0.13	0.07	0.15	0.15	0.09	0.67	1.36	3.09	2.26	0.84	3.47	5.05	0.07	0.04	0.01	0.20
ERR598996	0.22 µm	40m	17.7	North Atlantic Subtropical Gyral Province	1	7.31	0.44	1.85	0.84	1.17	0.16	0.18	0.10	0.06	0.03	0.03	0.00	0.09	0.04	0.18	2.02	0.46	0.02	0.06
ERR594294	0.22 µm	50m	16.8	South Atlantic Gyral Province	1	14.74	1.31	4.54	2.64	3.84	1.39	1.11	0.59	0.26	0.14	0.13	0.02	0.45	0.12	0.45	2.91	0.17	0.03	0.93
ERR599094	0.22 µm	50m	15.2	Mediterranean Sea, Black Sea Province	1	1.22	0.78	2.66	2.86	4.85	0.96	0.67	0.39	0.04	0.03	0.02	0.00	0.11	0.03	0.06	0.43	0.06	0.01	0.82
ERR598982	0.22 µm	30m	15	Benguela Current Coastal Province	1	3.18	0.87	1.27	3.41	3.84	1.03	0.86	0.70	0.07	0.04	0.03	0.00	0.11	0.04	0.14	14.21	0.31	0.01	0.58
ERR599001	0.22 µm	25m	14.3	North Atlantic Subtropical Gyral Province	1	1.04	0.40	0.60	1.18	1.79	1.37	0.88	0.59	0.06	0.04	0.04	0.00	0.11	0.04	0.20	22.98	0.19	0.02	0.48
ERR598942	0.22 µm	45m	13.2	North Pacific Subtropical and Polar Front Provinces	1	1.21	26.25	1.42	10.12	4.73	0.43	0.49	0.35	0.10	0.09	0.07	0.00	0.18	0.07	0.21	4.25	0.28	0.01	0.01
ERR599161	0.22 µm	120m	25.2	South Pacific Subtropical Gyre Province, North and South	1	2.60	0.49	0.91	0.50	0.49	0.22	0.21	0.13	0.11	0.07	0.07	0.01	0.16	0.07	0.24	0.27	0.14	0.02	1.00
ERR599100	0.22 µm	125m	24.9	Caribbean Province	1	2.07	0.70	1.03	0.86	0.81	0.32	0.45	0.19	0.25	0.19	0.18	0.03	0.34	0.16	0.50	0.44	0.26	0.04	0.09
ERR594342	0.22 µm	140m	23.7	South Pacific Subtropical Gyre Province, North and South	1	3.64	0.84	1.51	0.96	0.87	0.20	0.21	0.12	0.12	0.11	0.09	0.02	0.16	0.09	0.34	0.29	0.28	0.03	0.01
ERR598957	0.22 µm	155m	22.3	South Pacific Subtropical Gyre Province, North and South	1	1.24	0.76	0.92	0.80	0.71	0.30	0.40	0.19	0.14	0.10	0.08	0.01	0.20	0.07	0.18	0.11	0.08	0.01	0.00
ERR598972	0.22 µm	65m	22.3	Eastern Africa Coastal Province	1	5.16	0.44	1.21	0.60	0.59	0.13	0.14	0.10	0.13	0.11	0.11	0.01	0.14	0.10	0.44	0.73	0.29	0.04	0.08
HOTs	0.22 µm	75m	22	North Pacific Subtropical Gyre	5	1.77	0.12	0.46	0.21	0.24	0.01	0.04	0.07	0.02	0.00	0.00	0.00	0.09	0.01	0.13	0.07	0.05	0.01	0.01
ERR594298	0.22 µm	150m	21.6	South Atlantic Gyral Province	1	8.28	0.82	2.12	0.92	1.19	0.93	0.91	0.34	0.39	0.17	0.14	0.03	0.60	0.13	0.37	0.23	0.11	0.03	0.15
ERR598961	0.22 µm	90m	19.9	South Pacific Subtropical Gyre Province, North and South	1	4.49	1.37	2.74	1.40	1.53	0.46	0.48	0.23	0.16	0.10	0.09	0.01	0.24	0.09	0.28	0.68	0.10	0.02	0.00
ERR594336	0.22 µm	120m	19.3	South Atlantic Gyral Province	1	10.36	0.74	2.51	0.99	1.24	0.64	0.59	0.28	0.19	0.10	0.09	0.01	0.31	0.08	0.29	0.48	0.12	0.02	0.84
ERR599087	0.22 µm	60m	19	North Pacific Equatorial Countercurrent Province	1	0.74	1.37	0.72	1.04	0.95	0.28	0.31	0.24	0.12	0.08	0.06	0.02	0.21	0.06	0.27	0.90	0.27	0.02	0.13
ERR318618	0.22 µm	70m	19	Mediterranean Sea, Black Sea Province	1	1.03	0.12	1.27	0.17	0.24	0.03	0.04	0.03	0.01	0.01	0.01	0.00	0.04	0.01	0.00	0.02	0.00	0.00	0.01
ERR599073	0.22 µm	60m	18.4	Mediterranean Sea, Black Sea Province	1	17.10	1.85	8.27	5.17	9.29	0.94	0.78	0.40	0.12	0.08	0.06	0.01	0.20	0.07	0.21	0.62	0.15	0.01	0.02
ERR598986	0.22 µm	80m	16.8	North Atlantic Subtropical Gyral Province	1	8.30	0.78	2.26	1.90	2.82	1.75	1.24	0.69	0.57	0.09	0.07	0.01	0.37	0.07	0.29	7.31	0.51	0.03	0.87
ERR594315	0.22 µm	55m	16.2	Mediterranean Sea, Black Sea Province	1	1.73	2.46	4.30	6.15	10.49	11.41	7.25	4.27	0.14	0.11	0.09	0.02	0.49	0.09	0.30	7.21	0.35	0.03	1.38
ERR315859	0.22 µm	55m	15.7	Mediterranean Sea, Black Sea Province	1	1.50	1.37	4.28	5.20	9.33	2.52	1.62	0.95	0.08	0.05	0.05	0.01	0.21	0.06	0.13	1.23	0.22	0.02	0.23
MedDCM-SEP2013	0.22 µm	55m	15.5	Mediterranean Sea	2	1.02	1.31	1.93	4.69	8.73	0.52	0.42	0.29	0.06	0.04	0.04	0.00	0.11	0.04	0.13	7.05	0.63	0.02	0.09
Med-Io16-70mDCM	0.22 µm	70m	15.4	Mediterranean Sea	4	2.37	0.97	6.25	2.93	5.13	0.66	0.50	0.27	5.40	0.08	0.06	0.00	0.13	0.06	0.09	0.62	0.05	0.00	0.00
ERR598995	0.22 µm	115m	15.3	North Pacific Subtropical and Polar Front Provinces	1	1.39	2.15	1.35	2.49	3.07	4.31	2.91	1.72	0.32	0.15	0.13	0.02	0.57	0.12	0.36	1.78	0.21	0.02	0.11
MedDCM-SEP2014	0.22 µm	60m	15	Mediterranean Sea	3	3.32	1.60	16.47	5.31	9.61	0.70	0.62	0.36	0.09	0.07	0.05	0.00	0.17	0.06	0.11	0.86	0.06	0.01	0.03
Med-Ae1-75mDCM	0.22 µm	75m	15	Mediterranean Sea	4	0.61	0.43	0.76	1.05	1.74	0.91	0.67	0.36	4.05	0.06	0.04	0.02	0.21	0.04	0.11	0.27	0.05	0.01	0.10
Med-Io7-77mDCM	0.22 µm	77m	14.3	Mediterranean Sea	4	1.03	1.02	2.96	3.18	5.51	5.20	3.42	1.96	3.93	0.06	0.05	0.01	0.28	0.05	0.12	0.96	0.11	0.02	0.11
MedDCM-JUL2012	0.22 µm	75m	13.6	Mediterranean Sea	2	0.71	0.82	0.88	2.85	3.95	21.30	13.27	7.87	0.88	0.08	0.06	0.01	0.64	0.06	0.22	20.14	0.18	0.02	4.33
ERR598958	0.22 µm	250m	18.2	North Atlantic Subtropical Gyral Province	1	0.71	0.23	0.24	0.32	0.21	0.34	0.32	0.17	3.42	0.16	0.14	0.04	0.43	0.14	0.58	0.50	0.16	0.04	1.12
ERR599172	0.22 µm	270m	15.6	Indian Monsoon Gyres Province	1	0.04	0.05	0.03	0.07	0.02	0.05	0.05	0.02	0.45	1.83	3.18	2.32	0.22	2.90	2.31	0.01	0.01	0.02	0.05
ERR599109	0.22 µm	340m	15	Indian Monsoon Gyres Province	1	0.04	0.04	0.04	0.06	0.03	0.06	0.06	0.02	0.48	1.05	1.92	1.33	0.34	1.72	1.94	0.02	0.01	0.02	0.14
Med-Ae2-600mDeep	0.22 µm	600m	14	Mediterranean Sea	4	0.12	0.10	0.06	0.14	0.09	0.20	0.23	0.10	13.97	0.13	0.09	0.03	1.42	0.09	0.24	0.13	0.06	0.03	0.37
ERR598953	0.22 µm	177m	13	South Pacific Subtropical Gyre Province, North and South	1	0.24	0.33	0.21	0.39	0.24	0.49	0.55	0.29	1.55	0.52	1.10	0.76	5.07	1.27	2.08	0.14	0.05	0.02	1.23
ERR594290	0.22 µm	600m	12.1	Northwest Arabian Sea Upwelling Province	1	0.02	0.02	0.02	0.06	0.01	0.03	0.06	0.01	0.24	3.19	5.99	4.19	0.12	5.22	1.16	0.01	0.01	0.01	0.82
ERR599067	0.22 µm	380m	11.3	Chile-Peru Current Coastal Province	1	0.13	0.09	0.06	0.13	0.07	0.12	0.18	0.05	0.89	0.11	0.11	0.08	1.35	0.14	1.31	0.19	0.06	0.02	0.96
ERR599047	0.22 µm	640m	11	North Atlantic Subtropical Gyral Province	1	0.16	0.12	0.07	0.15	0.08	0.23	0.31	0.16	1.79	0.36	0.50	0.37	0.70	0.56	1.88	0.19	0.06	0.03	1.44
ERR599086	0.22 µm	350m	10.9	South Pacific Subtropical Gyre Province, North and South	1	0.21	0.16	0.12	0.20	0.11	0.23	0.28	0.13	1.36	0.39	0.30	0.28	2.65	0.43	4.51	0.25	0.08	0.03	1.12
ERR599020	0.22 µm	380m	10.3	South Pacific Subtropical Gyre Province, North and South	1	0.46	0.27	0.28	0.31	0.22	0.16	0.23	0.10	0.86	0.29	0.51	0.36	3.01	0.59	1.51	0.44	0.19	0.03	0.79
ERR598985	0.22 µm	640m	9.8	Caribbean Province	1	0.12	0.08	0.06	0.13	0.07	0.13	0.22	0.07	1.10	0.47	0.78	0.51	0.57	0.87	1.84	0.19	0.06	0.03	1.55
ERR599055	0.22 µm	480m	9.2	Pacific Equatorial Divergence Province	1	0.12	0.14	0.07	0.15	0.08	0.13	0.27	0.08	0.72	0.11	0.09	0.09	2.90	0.14	1.16	0.16	0.05	0.02	1.98
ERR599152	0.22 µm	375m	8.9	North Pacific Equatorial Countercurrent Province	1	0.08	0.08	0.05	0.10	0.05	0.09	0.12	0.04	0.71	2.26	5.31	3.95	0.34	6.02	8.04	0.04	0.01	0.02	0.53
ERR599004	0.22 µm	450m	8.2	North Pacific Equatorial Countercurrent Province	1	0.06	0.08	0.04	0.09	0.04	0.08	0.10	0.03	0.58	1.97	4.74	3.50	0.35	5.36	6.24	0.04	0.02	0.02	0.45
ERR594305	0.22 µm	600m	7.2	South Pacific Subtropical Gyre Province, North and South	1	0.33	0.07	0.10	0.10	0.07	0.08	0.25	0.04	0.57	0.33	0.34	0.29	0.29	0.44	4.12	0.21	0.07	0.02	3.45
ERR599166	0.22 µm	590m	5.1	Gulf Stream Province	1	0.18	0.18	0.12	0.27	0.12	0.42	0.61	0.32	1.76	1.43	3.03	2.16	4.85	3.22	1.51	0.20	0.08	0.03	2.36
ERR599115	0.22 µm	650m	4.9	North Pacific Subtropical and Polar Front Provinces	1	0.16	0.17	0.08	0.18	0.09	0.17	0.28	0.11	1.30	0.72	1.31	1.00	4.10	1.53	4.12	0.25	0.08	0.03	1.89
ERR598964	0.22 µm	740m	10.6	North Atlantic Subtropical Gyral Province	1	0.14	0.11	0.07	0.13	0.07	0.17	0.27	0.11	1.61	0.33	0.46	0.34	0.74	0.53	1.43	0.17	0.06	0.02	1.59
ERR598944	0.22 µm	800m	10.2	North Atlantic Subtropical Gyral Province	1	0.12	0.11	0.06	0.14	0.08	0.15	0.25	0.07	1.52	0.54	1.01	0.71	0.60	1.09	1.51	0.19	0.07	0.03	1.31
ERR598960	0.22 µm	850m	8.4	Eastern Africa Coastal Province	1	0.16	0.10	0.07	0.12	0.07	0.14	0.26	0.05	1.50	0.40	0.64	0.42	0.60	0.74	1.62	0.16	0.05	0.02	1.37
ERR599021	0.22 µm	1000m	7.7	Eastern Africa Coastal Province	1	0.14	0.09	0.06	0.13	0.06	0.14	0.25	0.06	1.29	0.77	1.48	1.04	0.63	1.63	2.41	0.14	0.05	0.03	1.41
ERR598947	0.22 µm	700m	7	South Atlantic Gyral Province	1	0.18	0.16	0.10	0.20	0.10	0.32	0.50	0.23	1.95	0.46	0.88	0.61	3.75	1.00	1.06	0.20	0.06	0.03	2.13
ERR599124	0.22 µm	800m	5.9	South Atlantic Gyral Province	1	0.11	0.09	0.05	0.12	0.05	0.13	0.28	0.08	0.84	0.91	2.11	1.44	1.16	2.38	0.93	0.08	0.04	0.03	1.51
ERR599000	0.22 µm	800m	4.7	South Atlantic Gyral Province	1	0.07	0.06	0.03	0.09	0.04	0.06	0.17	0.03	0.67	1.32	2.90	2.07	0.52	3.22	2.23	0.09	0.03	0.03	1.30
ERR599048	0.22 µm	800m	4.7	South Atlantic Gyral Province	1	2.29	0.19	0.54	0.25	0.24	0.07	0.10	0.04	0.32	0.30	0.52	0.47	0.20	0.61	1.40	0.11	0.05	0.02	0.50
ERR599149	0.22 µm	800m	4.2	South Atlantic Gyral Province	1	0.11	0.07	0.05	0.09	0.04	0.07	0.15	0.03	0.62	0.81	1.54	1.27	0.43	1.66	2.64	0.10	0.04	0.02	0.97
ERR599125	0.22 µm	790m	0.5	Antarctic Province	1	0.14	0.14	0.10	0.19	0.08	0.31	0.45	0.18	1.16	2.98	6.96	5.14	3.50	5.76	1.09	0.15	0.06	0.03	1.18
KM3	0.22 µm	3000m	14	Mediterranean Sea	6	0.00	0.00	0.00	0.00	0.28	0.00	0.00	0.00	0.07	6.51	1.37	1.63	0.13	1.23	0.29	0.28	0.00	0.00	0.19
Med-Io17-3500mDeep	0.22 µm	3500m	14	Mediterranean Sea	4	0.03	0.04	0.03	0.05	0.02	0.05	0.05	0.04	0.14	72.89	29.25	20.95	0.08	23.92	3.35	0.02	0.03	0.03	0.14
SRR488330	0.22 µm	2000m	3	Guaymas Basin	7	0.06	0.01	0.02	0.06	0.01	0.11	0.15	0.00	0.05	0.14	0.55	0.38	2.72	0.84	0.17	0.10	0.06	0.03	2.24
SRR488331	0.22 µm	2000m	3	Guaymas Basin	7	0.05	0.07	0.07	0.21	0.01	0.19	0.16	0.17	0.17	0.17	0.72	0.27	3.06	1.18	0.28	0.15	0.05	0.04	2.44
SRR2046235	0.22 µm	2040m	3	Cayman	7	0.06	0.07	0.04	0.09	0.04	0.09	0.12	0.02	3.89	2.44	6.38	5.01	0.20	4.82	1.40	0.03	0.03	0.02	0.22
SRR2046236	0.22 µm	2238m	3	Cayman	7	0.06	0.06	0.03	0.06	0.03	0.06	0.08	0.01	5.18	3.04	8.04	6.66	0.20	5.95	1.37	0.04	0.02	0.02	0.23
SRR2046222	0.22 µm	2327m	3	Cayman	7	0.07	0.07	0.04	0.09	0.04	0.10	0.11	0.02	0.74	4.41	11.36	9.04	0.24	8.53	1.69	0.05	0.03	0.02	0.23
SRR2046238	0.22 µm	4869m	1	Cayman	7	0.05	0.04	0.03	0.07	0.02	0.06	0.09	0.02	0.41	8.00	22.60	19.88	0.18	14.85	1.86	0.05	0.04	0.03	0.74
SRR2046237	0.22 µm	4946m	1	Cayman	7	0.07	0.07	0.05	0.08	0.04	0.08	0.07	0.03	0.52	8.62	24.90	21.64	0.19	16.47	1.89	0.06	0.03	0.02	0.20
SRR2046221	0.22 µm	4950m	1	Cayman	7	0.10	0.07	0.05	0.11	0.04	0.08	0.13	0.04	4.77	12.39	36.11	31.20	0.27	23.89	2.91	0.08	0.05	0.03	0.23

(1) Sunagawa et al. (2015); (2) Martin-Cuadrado et al. (2015); (3) This work; (4) Mizuno et al. (2016); (5) Coleman and Chisholm, (2010); (6) Martin-Cuadrado et al. (2007); (7) Li et al. (2015)

Highlighted in grey, list of the 33 metagenomes that were assembled and used for the differential coverage in the binning procedure.

Supplementary Table 3. List of contig of the MG-III composite genomes.

CG-EPI-1			CG-EPI-2			CG-EPI-3			CG-EPI-4			CG-EPI-5			CG-EPI-6			CG-BATHY-1			CG-BATHY-2		
Contig	GC	length (bp)	Contig	GC	length (bp)	Contig	GC	length (bp)	Contig	GC	length (bp)	Contig	GC	length (bp)	Contig	GC	length (bp)	Contig	GC	length (bp)	Contig	GC	length (bp)
CG-Epi1-C1	34.8	135963	CG-Epi2-C1	35.8	101994	CG-Epi3-C1	34.8	71763	CG-Epi4-C1	36.7	80064	CG-Epi5-C1	35.8	47627	CG-Epi6-C1	36.8	50314	CG-Bathy1-C1	36.6	211731	CG-Bathy2-C1	63.5	130682
CG-Epi1-C2	36.8	98555	CG-Epi2-C2	35.2	92741	CG-Epi3-C2	35.6	63159	CG-Epi4-C2	36.1	59597	CG-Epi5-C2	37.1	26037	CG-Epi6-C2	35.8	41924	CG-Bathy1-C2	36.7	133274	CG-Bathy2-C2	63.5	85590
CG-Epi1-C3	36.6	91163	CG-Epi2-C3	36.1	78522	CG-Epi3-C3	35.8	49058	CG-Epi4-C3	36.8	59301	CG-Epi5-C3	36.7	25001	CG-Epi6-C3	37.0	41453	CG-Bathy1-C3	36.5	91842	CG-Bathy2-C3	67.2	63709
CG-Epi1-C4	36.4	78844	CG-Epi2-C4	37.4	65383	CG-Epi3-C4	35.5	47213	CG-Epi4-C4	37.0	47475	CG-Epi5-C4	36.5	22352	CG-Epi6-C4	36.2	39969	CG-Bathy1-C4	37.0	77586	CG-Bathy2-C4	64.2	63432
CG-Epi1-C5	36.8	73433	CG-Epi2-C5	35.5	63969	CG-Epi3-C5	35.8	41384	CG-Epi4-C5	36.3	40977	CG-Epi5-C5	36.8	21574	CG-Epi6-C5	36.0	37884	CG-Bathy1-C5	38.7	68229	CG-Bathy2-C5	62.6	62484
CG-Epi1-C6	37.9	68690	CG-Epi2-C6	36.1	51772	CG-Epi3-C6	35.6	34927	CG-Epi4-C6	34.8	37978	CG-Epi5-C6	36.3	18166	CG-Epi6-C6	33.7	25846	CG-Bathy1-C6	38.1	65587	CG-Bathy2-C6	62.3	62026
CG-Epi1-C7	37.1	58515	CG-Epi2-C7	36.3	48636	CG-Epi3-C7	35.9	33678	CG-Epi4-C7	36.0	37801	CG-Epi5-C7	36.4	18034	CG-Epi6-C7	36.8	25634	CG-Bathy1-C7	43.6	43732	CG-Bathy2-C7	66.1	47436
CG-Epi1-C8	36.8	57181	CG-Epi2-C8	36.3	41423	CG-Epi3-C8	36.2	32361	CG-Epi4-C8	36.6	34409	CG-Epi5-C8	35.7	16719	CG-Epi6-C8	36.2	22443	CG-Bathy1-C8	38.3	41508	CG-Bathy2-C8	64.7	37420
CG-Epi1-C9	36.7	53175	CG-Epi2-C9	36.6	39937	CG-Epi3-C9	35.2	28243	CG-Epi4-C9	36.7	31390	CG-Epi5-C9	35.9	16426	CG-Epi6-C9	36.6	21880	CG-Bathy1-C9	37.9	41315	CG-Bathy2-C9	65.0	33792
CG-Epi1-C10	37.2	50091	CG-Epi2-C10	35.7	38179	CG-Epi3-C10	36.6	28004	CG-Epi4-C10	36.7	31342	CG-Epi5-C10	35.4	16368	CG-Epi6-C10	36.2	21094	CG-Bathy1-C10	36.0	37030	CG-Bathy2-C10	62.7	27467
CG-Epi1-C11	36.3	40290	CG-Epi2-C11	36.5	37752	CG-Epi3-C11	36.9	27182	CG-Epi4-C11	36.8	27609	CG-Epi5-C11	36.2	15933	CG-Epi6-C11	35.1	20996	CG-Bathy1-C11	43.0	35415	CG-Bathy2-C11	64.4	26425
CG-Epi1-C12	36.8	37293	CG-Epi2-C12	37.1	31968	CG-Epi3-C12	36.5	26391	CG-Epi4-C12	35.4	20476	CG-Epi5-C12	34.9	15869	CG-Epi6-C12	36.9	20201	CG-Bathy1-C12	37.4	27476	CG-Bathy2-C12	64.8	25944
CG-Epi1-C13	37.2	36578	CG-Epi2-C13	36.2	31324	CG-Epi3-C13	35.9	25790	CG-Epi4-C13	37.2	20393	CG-Epi5-C13	36.2	15690	CG-Epi6-C13	36.4	19776	CG-Bathy1-C13	36.5	23386	CG-Bathy2-C13	65.8	25194
CG-Epi1-C14	37.2	34349	CG-Epi2-C14	36.2	30879	CG-Epi3-C14	34.3	25074	CG-Epi4-C14	36.6	19527	CG-Epi5-C14	36.6	14852	CG-Epi6-C14	36.1	18182	CG-Bathy1-C14	36.2	22942	CG-Bathy2-C14	66.2	24873
CG-Epi1-C15	36.4	33957	CG-Epi2-C15	37.4	27764	CG-Epi3-C15	36.7	24609	CG-Epi4-C15	36.1	18526	CG-Epi5-C15	38.1	13655	CG-Epi6-C15	36.4	14526	CG-Bathy1-C15	38.0	22193	CG-Bathy2-C15	66.1	21939
CG-Epi1-C16	36.1	29648	CG-Epi2-C16	33.8	26756	CG-Epi3-C16	35.4	20644	CG-Epi4-C16	36.6	17553	CG-Epi5-C16	35.8	13629	CG-Epi6-C16	37.3	13654	CG-Bathy1-C16	36.7	21128	CG-Bathy2-C16	64.7	12380
CG-Epi1-C17	35.8	27053	CG-Epi2-C17	36.3	23686	CG-Epi3-C17	37.3	20166	CG-Epi4-C17	36.8	17247	CG-Epi5-C17	37.0	13467	CG-Epi6-C17	37.1	12311	CG-Bathy1-C17	35.3	18158	CG-Bathy2-C17	65.9	12098
CG-Epi1-C18	37.3	26563	CG-Epi2-C18	37.1	22912	CG-Epi3-C18	36.1	20159	CG-Epi4-C18	37.0	16984	CG-Epi5-C18	35.6	12483	CG-Epi6-C18	34.7	11537	CG-Bathy1-C18	37.1	17274	CG-Bathy2-C18	67.5	10866
CG-Epi1-C19	37.4	26142	CG-Epi2-C19	34.9	19544	CG-Epi3-C19	34.0	19853	CG-Epi4-C19	34.8	16641	CG-Epi5-C19	36.8	12156	CG-Epi6-C19	36.0	11219	CG-Bathy1-C19	36.7	11671			
CG-Epi1-C20	36.8	23896	CG-Epi2-C20	36.8	19543	CG-Epi3-C20	36.4	18732	CG-Epi4-C20	36.5	16164	CG-Epi5-C20	34.7	11809	CG-Epi6-C20	36.5	11161	CG-Bathy1-C20	38.8	10952			
CG-Epi1-C21	37.5	23263	CG-Epi2-C21	37.6	19118	CG-Epi3-C21	37.0	17297	CG-Epi4-C21	37.0	15964	CG-Epi5-C21	35.7	11221	CG-Epi6-C21	39.7	10474	CG-Bathy1-C21	36.0	10381			
CG-Epi1-C22	37.6	21975	CG-Epi2-C22	36.3	18555	CG-Epi3-C22	39.9	15409	CG-Epi4-C22	36.7	14428	CG-Epi5-C22	37.5	10057	CG-Epi6-C22	35.9	10049	CG-Bathy1-C22	37.6	10120			
CG-Epi1-C23	35.8	19014	CG-Epi2-C23	36.4	17861	CG-Epi3-C23	37.5	14390	CG-Epi4-C23	37.6	12998				CG-Epi6-C23	35.5	9955						
CG-Epi1-C24	35.6	17494	CG-Epi2-C24	36.7	17829	CG-Epi3-C24	36.6	13412	CG-Epi4-C24	36.1	12867				CG-Epi6-C24	38.0	9081						
CG-Epi1-C25	35.2	14922	CG-Epi2-C25	35.3	17776	CG-Epi3-C25	37.1	13266							CG-Epi6-C25	34.1	8023						
			CG-Epi2-C26	35.5	17562	CG-Epi3-C26	38.7	13151							CG-Epi6-C26	36.9	5465						
			CG-Epi2-C27	36.2	16729	CG-Epi3-C27	35.5	12130															
			CG-Epi2-C28	37.2	16047	CG-Epi3-C28	35.3	12129															
			CG-Epi2-C29	34.8	15419	CG-Epi3-C29	35.0	10830															
			CG-Epi2-C30	36.8	14321	CG-Epi3-C30	36.2	10429															
			CG-Epi2-C31	36.5	14306																		
			CG-Epi2-C32	36.9	13761																		
			CG-Epi2-C33	34.3	12919																		
			CG-Epi2-C34	35.0	12691																		
			CG-Epi2-C35	35.9	12415																		
			CG-Epi2-C36	34.4	12304																		
			CG-Epi2-C37	35.9	12225																		
			CG-Epi2-C38	36.5	12201																		
			CG-Epi2-C39	35.9	12194																		
			CG-Epi2-C40	36.2	11337																		
			CG-Epi2-C41	33.9	10850																		
			CG-Epi2-C42	36.3	10816																		
			CG-Epi2-C43	38.4	10309																		

Supplementary Table 4. Housekeeping genes found in the MG-III bins and the CG-MG-III bins (as Naransigarao et al. 2012).

			Number of housekeeping genes found in the MG-III bins										Number of housekeeping genes found in the CG-MG-III bins								Number of housekeeping genes found in reference genomes			
			Epi1	Epi2A	Epi2B	Epi2C	Epi3	Epi4	Epi5	Epi6	Bathy1	Bathy2	Epi1	Epi2	Epi3	Epi4	Epi5	Epi6	Bathy1	Bathy2	Cayman 92	Cayman 93	Guaymas 31	Guaymas 32
		% Completness estimation (Raes et al. 2007)	**80**	**54.3**	**57.1**	**5.7**	**48.6**	**34.3**	**2.9**	**51.4**	**54.3**	**68.6**	**85.7**	**80**	**62.9**	**34.3**	**2.9**	**57.1**	**60**	**68.6**	**71.4**	**34.3**	**91.4**	**94.3**
		% Completness estimation (Albertsen et al. 2013)	**35.1**	**16.2**	**18.0**	**4.5**	**19.8**	**15.3**	**3.6**	**15.3**	**23.4**	**18.9**	**34.2**	**30.6**	**24.3**	**16.2**	**4.5**	**21.6**	**26.1**	**21.6**	**21.6**	**15.3**	**33.3**	**37.8**
		% Completness estimation (Narasingarao et al. 2012)	**83.0**	**45.3**	**54.7**	**5.7**	**45.3**	**39.6**	**7.6**	**54.7**	**60.4**	**58.5**	**84.9**	**75.2**	**62.3**	**43.4**	**9.4**	**60.4**	**64.2**	**58.5**	**67.9**	**35.9**	**92.5**	**90.6**
		# of genome-species inside the bin	**3.07**	**1.08**	**1.03**	**1.00**	**1.00**	**1.00**	**1.00**	**1.48**	**1.09**	**2.32**	**1.18**	**1.43**	**1**	**1**	**1**	**1.0385**	**1.14**	**1**	**1**	**1**	**1**	**1.86**
Ribosomal genes		16S RNA	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	1	-	-	-	-	3
		23S RNA	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	1	-	-	-	-	3
COG	**COG classification**	**Function**																						
	COG0008	GlnS COG0008 472 Glutamyl- and glutaminyl-tRNA synthetases	3	1	1	-	-	-	-	-	1	-	2	2	-	1	1	-	1	-	1	-	1	1
	COG0012	Predicted GTPase, probable translation factor	1	-	-	-	-	-	-	-	1	-	1	-	-		-	-	2	-		-	1	1
	COG0013	AlaS COG0013 879 Alanyl-tRNA synthetase	1	1	-	-	1	1	-	-	2	1	1	2	1	1	-	-	1	1		-	1	1
	COG0016	PheS COG0016 335 Phenylalanyl-tRNA synthetase alpha subunit	3	1	-	1	1	-	-	-	1	1	1	2	1	-	-	-	1	1	1	-	1	2
	COG0018	ArgS COG0018 577 Arginyl-tRNA synthetase	2	-	-	-	1	1	1	1	1	-	2	-	1	1	1	1	1	-	1	1	1	2
	COG0024	Map COG0024 255 Methionine aminopeptidase	5	1	1	-	1	1	1	1	2	-	1	2	1	1	1	1	3	-	1	-	1	2
	COG0048	RpsL COG0048 129 Ribosomal protein S12	2	-	1	-	-	1	-	2	1	-	1	1	1	1	-	1	1	-	1	-	1	2
	COG0049	RpsG COG0049 148 Ribosomal protein S7	2	-	1	-	-	1	-	2	1	-	1	1	1	1	-	1	1	-	1	-	1	2
	COG0051	RpsJ COG0051 104 Ribosomal protein S10	1	-	1	-	-	-	-	2	1	-	1	1	-	-	-	1	1	-	1	-	1	2
	COG0060	IleS COG0060 933 Isoleucyl-tRNA synthetase	2	-	-	-	-	-	-	1	-	-	-	1	-	-	-	1	-	-		-	1	2
	COG0071	IbpA COG0071 146 Molecular chaperone (small heat shock protein)	2	-	-	-	-	1	-	1	1	2	1	-	-	1	-	1	1	1	1	-	1	2
	COG0080	RplK COG0080 141 Ribosomal protein L11	2	1	-	-	-	1	-	1	1	-	1	2	1	1	-	1	1		1	-	1	2
	COG0081	RplA COG0081 228 Ribosomal protein L1	1	-	1	-	-	-	-	-	1	1	1	1	-	-	-	-	1	1	1	1	1	1
	COG0085	RpoB COG0085 1060 DNA-directed RNA polymerase, beta subunit/140 kD subunit	-	-	1	-	1	1	-	3	1	3	-	1	1	1	-	1	1	1	1	-	1	2
	COG0086	RpoC COG0086 808 DNA-directed RNA polymerase, beta subunit/160 kD subunit	-	-	1	-	1	1	-	3	1	3	-	1	1	1	-	1	1	1	1	1	1	1
	COG0087	RplC COG0087 218 Ribosomal protein L3	5	-	1	-	-	-	-	-	-	3	1	1	1	-	-	-	-	1	1	1	1	2
	COG0088	RplD COG0088 214 Ribosomal protein L4	5	-	1	-	1	-	-	-	-	3	1	1	1	-	-	-	-	1	1	1	1	2
	COG0090	RplB COG0090 275 Ribosomal protein L2	5	1	1	-	1	-	-	-	-	3	1	1	1	-	-	-	-	1	1	1	1	1
	COG0091	RplV COG0091 120 Ribosomal protein L22	5	1	1	-	1	-	-	-	-	3	1	1	1	-	-	-	-	1	1	1	1	2
	COG0092	RpsC COG0092 233 Ribosomal protein S3	5	1	1	-	1	-	-	-	-	3	1	1	1	-	-	-	-	1	1	1	1	2
	COG0093	RplN COG0093 122 Ribosomal protein L14	6	1	1	-	1	-	-	1	-	3	1	2	1	-	-	1	-	1	1	1	1	2
	COG0094	RplE COG0094 180 Ribosomal protein L5	6	1	1	-	1	-	-	1	1	3	1	2	1	-	-	1	1	1	1	-	1	2
	COG0096	RpsH COG0096 132 Ribosomal protein S8	6	1	1	-	1	-	-	1	1	3	1	2	1	-	-	1	1	1	1	-	1	2
	COG0097	RplF COG0097 178 Ribosomal protein L6P/L9E	6	1	1	-	1	-	-	1	1	3	1	2	1	-	-	1	1	1	1	-	1	2
	COG0098	RpsE COG0098 181 Ribosomal protein S5	6	1	1	-	1	-	-	1	1	3	1	2	1	-	-	1	1	1	1	-	1	2
	COG0099	RpsM COG0099 121 Ribosomal protein S13	2	-	-	-	-	1	-	-	-	1	2	-	-	1	-	-	-	1	-	-	1	2
	COG0100	RpsK COG0100 129 Ribosomal protein S11	2	-	-	-	-	1	-	-	-	2	2	-	-	1	-	-	-	1	1	1	1	2
	COG0102	RplM COG0102 148 Ribosomal protein L13	1	1	-	-	-	1	-	1	-	3	1	1	-	1	-	1	1	1	-	-	1	2
	COG0103	RpsI COG0103 130 Ribosomal protein S9	1	1	-	-	-	1	-	1	-	3	1	1	-	1	-	1	1	1	-	-	1	2
	COG0124	HisS COG0124 429 Histidyl-tRNA synthetase	-	-	-	-	-	-	-	2	1	-	1	-	-	-	-	-	1	-	1	-	1	2
	COG0130	TruB COG0130 271 Pseudouridine synthase	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	1	-	-	1	1	2
	COG0143	MetG COG0143 558 Methionyl-tRNA synthetase	2	1	1	-	-	-	-	1	-	2	1	-	1	-	-	1	-	1	1	-	1	1
	COG0164	RnhB COG0164 199 Ribonuclease HII	2	2	-	1	1	-	-	-	2	-	1	1	1	-	-	-	2	-	-	-	1	2
	COG0177	Nth COG0177 211 Predicted EndoIII-related endonuclease	1	-	-	1	1	1	-	-	1	1	1	1	1	1	-	-	1	1	1	-	1	2
	COG0180	TrpS COG0180 314 Tryptophanyl-tRNA synthetase	1	-	1	-	1	1	1	1	1	-	3	1	1	1	1	1	1	-	1	1	1	1
	COG0186	RpsQ COG0186 87 Ribosomal protein S17	6	1	1	-	1	-	-	1	-	3	1	2	1	-	-	1	-	1	1	1	1	2
	COG0197	RplP COG0197 146 Ribosomal protein L16/L10E	5	2	-	-	1	1	-	-	1	-	1	1	1	1	-	-	1	-	-	-	1	2
	COG0200	RplO COG0200 152 Ribosomal protein L15	6	1	1	-	1	-	-	1	1	3	1	2	1	-	-	-	1	1	1	-	1	2
	COG0201	SecY COG0201 436 Preprotein translocase subunit SecY	5	1	1	-	-	-	-	1	1	3	1	2	-	-	-	-	1	1	1	-	1	2
	COG0250	NusG COG0250 178 Transcription antiterminator	1	1	-	-	-	1	-	1	-	-	1	2	1	1	-	1	1	-	-	-	1	2
	COG0256	RplR COG0256 125 Ribosomal protein L18	6	1	1	-	1	-	-	1	1	3	1	2	1	-	-	-	1	1	1	-	1	1
	COG0441	ThrS COG0441 589 Threonyl-tRNA synthetase	1	-	1	-	-	-	-	3	1	-	1	1	-	-	-	1	1	-	1	1	1	2
	COG0459	GroL COG0459 524 Chaperonin GroEL (HSP60 family)	1	-	-	-	1	1	-	1	1	1	2	1	1	1	-	1	2	1	1	1	2	4
	COG0468	RecA COG0468 279 RecA/RadA recombinase	4	-	2	-	1	1	1	4	1	1	1	2	1	1	1	2	1	1	1	1	1	4
	COG0480	FusA COG0480 697 Translation elongation factors (GTPases)	2	-	1	-	-	1	-	2	1	-	1	1	1	1	-	1	1	-	1	-	1	1
	COG0495	LeuS COG0495 814 Leucyl-tRNA synthetase	1	1	-	-	-	-	-	-	-	-	1	1	-	-	-	-	-	-	1	1	1	2
	COG0522	RpsD COG0522 205 Ribosomal protein S4 and related proteins	2	-	-	-	-	1	-	-	-	1	2	-	-	1	-	-	-	1	1	-	1	2
	COG0525	ValS COG0525 877 Valyl-tRNA synthetase	-	-	1	-	-	-	-	-	-	3	1	1	1	-	-	-	-	1	-	1	1	2
	COG0532	InfB COG0532 509 Translation initiation factor 2 (IF-2; GTPase)	1	-	1	-	-	-	-	-	1	1	1	2	1	1	-	1	1	1	-	1	1	1

Supplementary Table 5. Protein categories based on arCOG database.

									Epipelagic CG-MGIII bins						Pelagic CG-MGIII bins		Other Euryarchaea		
		arCOG clasification	gene	product	COG classification	pFAM domain	cdc	TIGR classification	Epi1	Epi2	Epi3	Epi4	Epi5	Epi6	Bathy1	Bathy-2	A.boneii T469	MG2-GG3	Thalassoarchaea
INFORMATION STORAGE AND PROCESSING																			
arCOG00415	L	Replication, recombination and repair	RecA	RecA/RadA recombinase	COG00468	pfam14520,pfam0842	cd01123	TIGR02236	1	1	1	1	1	1	1	1	1	1	1
arCOG04143	L	Replication, recombination and repair	-	DNA topoisomerase VI, subunit A	COG01697	pfam04406	cd00223		1	1	1	1	1	1	1	1	1	1	5
arCOG04455	L	Replication, recombination and repair	HYS2	Archaeal DNA polymerase II, small subunit/DNA polymerase delta, subunit B	COG01311	pfam01336,pfam0404	cd04490,cd07386		1	1	1	1		1	1	1	1	1	1
arCOG00459	L	Replication, recombination and repair	Nth	EndoIII-related endonuclease	COG00177	pfam00730,pfam1057	cd00056	TIGR01083	1	1	1	1			1	1	2	1	1
arCOG00417	L	Replication, recombination and repair	RecA	RecA/RadA recombinase	COG00468		cd01394	TIGR02237	1	1	1		1	2	1	1	1		
arCOG04367	L	Replication, recombination and repair	GyrA	Type IIA topoisomerase (DNA gyrase/topo II, topoisomerase IV), A subunit	COG00188	pfam00521,pfam0398	cd00187	TIGR01063	1	2	1	1	1	1	1	1	1	1	3
arCOG00872	L	Replication, recombination and repair	MPH1	ERCC4-like helicase	COG01111	pfam00270,pfam0027	cd00046,cd120	TIGR00643,TIGR	1	2	1	1		1	1	1	1	1	2
arCOG04447	L	Replication, recombination and repair	-	Archaeal DNA polymerase II, large subunit	COG01933	pfam03833,pfam0984	5	TIGR00354	1	2		1			1	1	1	1	1
arCOG00551	L	Replication, recombination and repair	-	DNA replication initiation complex subunit, GINS15 family	COG01711		cd11714		1		1				1	1	1		
arCOG04990	L	Replication, recombination and repair	-	Predicted endonuclease, contains HTH and PD-DExK nuclease domains		pfam14811			1			1			1	1		1	2
arCOG00467	L	Replication, recombination and repair	CDC6	Cdc6-related protein, AAA superfamily ATPase	COG01474	pfam13401,pfam0907	cd00009,cd087	TIGR02928	2	1	1	1		1	1	1	1	2	6
arCOG02610	L	Replication, recombination and repair	ScpA	Chromatin segregation and condensation protein Rec8/ScpA/Scc1, kleisin fam	COG01354	pfam02616				1				1	1	1	1	1	2
arCOG00439	L	Replication, recombination and repair	MCM2	Predicted ATPase involved in replication control, Cdc46/Mcm family	COG01241	pfam14551,pfam0049	cd00009				1	1	1	1	1	1	1	1	2
arCOG00427	L	Replication, recombination and repair	RecJ/Cdc45	Single-stranded DNA-specific exonuclease RecJ	COG00608										1	1	2		
arCOG02258	L	Replication, recombination and repair	-	RPA family protein, a subunit of RPA complex in P.furiosus	COG03390				1	1	1	1	1	1	1		1	1	3
arCOG04582	L	Replication, recombination and repair	Dpo4/DinP	Family Y DNA polymerase	COG00389	pfam00817,pfam1179	cd03586		1	1	1	1	1		1			1	1
arCOG01166	L	Replication, recombination and repair	MutL	DNA mismatch repair enzyme (predicted ATPase)	COG00323	pfam13589,pfam0111	cd00075,cd007	TIGR00585	1	1	1	1		1	1		1	1	2
arCOG01165	L	Replication, recombination and repair	-	DNA topoisomerase VI, subunit B	COG01389	pfam02518,pfam0348	cd00075,cd008	TIGR01052	1	1	1	1			1		1	1	4
arCOG01486	L	Replication, recombination and repair	RnmV	5S rRNA maturation endonuclease (Ribonuclease M5), contains TOPRIM doma	COG01658	pfam01751	cd01027		1	1	1		1		1		1	1	1
arCOG04121	L	Replication, recombination and repair	RnhB	Ribonuclease HII	COG00164	pfam01351	cd07180	TIGR00729	1	1	1				1		1	1	2
arCOG00470	L	Replication, recombination and repair	HolB	ATPase involved in DNA replication HolB, large subunit	COG00470	pfam00004	cd00009	TIGR02397	1	2	1	1	1		1		1	1	1
arCOG04371	L	Replication, recombination and repair	GyrB	Type IIA topoisomerase (DNA gyrase/topo II, topoisomerase IV), B subunit	COG00187	pfam02518,pfam0020	cd00075,cd008	TIGR01059	1	2	1	1			1		1	1	3
arCOG04110	L	Replication, recombination and repair	PRI1	Eukaryotic-type DNA primase, catalytic (small) subunit	COG01467	pfam01896	cd04860	TIGR00335	1	2	2			1	1		1	2	5
arCOG00488	L	Replication, recombination and repair	DnaN	DNA polymerase sliding clamp subunit (PCNA homolog)	COG00592	pfam00705,pfam0274	cd00577	TIGR00590	1		1	1	1	1	1		1	1	2
arCOG01078	L	Replication, recombination and repair	-	NUDIX family hydrolase	COG00494	pfam00293	cd03428			1	1				1		2	1	1
arCOG00787	L	Replication, recombination and repair	-	UvrD/Rep family helicase fused to exonuclease family domain	COG02887	pfam12705	cd09637	TIGR01249,TIGR00372		1		1			1		1		
arCOG01527	L	Replication, recombination and repair	TopA	Topoisomerase IA	COG00550	pfam01751,pfam0113	cd03362,cd001	TIGR01057							1		1	1	2
arCOG04050	L	Replication, recombination and repair	FEN1	5'-3' exonuclease	COG00258	pfam00752,pfam0086	cd09867	TIGR03674							1		1	1	2
arCOG00558	L	Replication, recombination and repair	SrmB	Superfamily II DNA and RNA helicase	COG00513	pfam00270,pfam0027	cd00268,cd000	TIGR01389							1		1	3	3
arCOG00129	L	Replication, recombination and repair	-	DNA modification methylase	COG00863	pfam01555,pfam01555		TIGR01177							1		1		
arCOG00328	L	Replication, recombination and repair	PolB3	DNA polymerase PolB3	COG00417	pfam03104,pfam0013	cd05781,cd055	TIGR00592	1	1					2	1	2	1	4
arCOG00905	L	Replication, recombination and repair	-	Uracil-DNA glycosylase	COG01573	pfam03167	cd10031	TIGR00758		1					2	1	2		
arCOG01510	L	Replication, recombination and repair	RPA1	Single-stranded DNA-binding replication protein A (RPA), large (70 kD) subuni	COG01599		cd04491,cd04491		2	3	2	2	1	2	2	2	3	2	3
arCOG01072	L	Replication, recombination and repair	-	NUDIX family hydrolase	COG00494	pfam00293	cd03426		1	1	1	1		1	2				
arCOG03013	L	Replication, recombination and repair	PRI2	Eukaryotic-type DNA primase, large subunit	COG02219	pfam04104	cd06560		1	1	1				2		1	1	1
arCOG00469	L	Replication, recombination and repair	HolB	ATPase involved in DNA replication HolB, small subunit	COG00470	pfam13177,pfam0854	cd00009	TIGR02397	1	1	1				2		1	2	2
arCOG00553	L	Replication, recombination and repair	BRR2	Replicative superfamily II helicase	COG01204	pfam00270,pfam0027	cd00046,cd000	TIGR04121	2	3	1	2	1		3		1	1	4
arCOG01894	L	Replication, recombination and repair	Nfo	Endonuclease IV	COG00648	pfam01261	cd00019	TIGR00587	1	1	1		1			1	1	1	4
arCOG01526	L	Replication, recombination and repair	TopG2	Reverse gyrase	COG01110	pfam00270,pfam0175	cd00046,cd033	TIGR01054	1	2	1	1				1	1		2
arCOG00464	L	Replication, recombination and repair	AlkA	3-methyladenine DNA glycosylase/8-oxoguanine DNA glycosylase	COG00122	pfam07934,pfam0073	cd00056	TIGR00588	1		1			1		1	1	1	
arCOG03142	L	Replication, recombination and repair	-	Nuclease of RNase H fold, RuvC/YqgF family					1		1					1	1	1	1
arCOG00368	L	Replication, recombination and repair	SbcC	ATPase involved in DNA repair, SbcC	COG00419	pfam13476,pfam1351	cd03240,cd001	TIGR00611,TIGR02168		1				1		1	3		
arCOG02897	L	Replication, recombination and repair	MutS	Mismatch repair ATPase (MutS family)	COG00249	pfam01624,pfam0518	cd03284	TIGR01070		1				2		1	1	1	1
arCOG02724	L	Replication, recombination and repair	Ada	Methylated DNA-protein cysteine methyltransferase	COG00350	pfam01035	cd06445	TIGR00589		1						1	1	1	1
arCOG00397	L	Replication, recombination and repair	SbcD	DNA repair exonuclease, SbcD	COG00420	pfam00149	cd00840									1	2		
arCOG02257	L	Replication, recombination and repair	-	RPA family protein, a subunit of RPA complex in P.furiosus	COG03390											1			
arCOG08649	L	Replication, recombination and repair	-	Topoisomerase IB	COG03569	pfam02919,pfam0102	cd00660,cd00659		1	1		1						1	
arCOG04281	L	Replication, recombination and repair	DnaG	DNA primase (bacterial type)	COG00358	pfam13662	cd01029	TIGR01391	1	2							1	1	2
arCOG01073	L	Replication, recombination and repair	-	NUDIX family hydrolase	COG00494	pfam00293	cd03424	TIGR00052	1			1					1		
arCOG01347	L	Replication, recombination and repair	CDC9	ATP-dependent DNA ligase	COG01793	pfam04675,pfam0106	cd07901,cd079	TIGR00574	1								1		
arCOG02840	L	Replication, recombination and repair	PhrB	Deoxyribodipyrimidine photolyase	COG00415	pfam00875,pfam0344	1	TIGR03556	2	3	1							2	7
arCOG05109	L	Replication, recombination and repair	DnaQ	DNA polymerase III, epsilon subunit or related 3'-5' exonuclease	COG00847	pfam00929	cd06127	TIGR00573		1								1	
arCOG00890	L	Replication, recombination and repair	-	DNA modification methylase	COG00863	pfam01555				1									
arCOG01898	L	Replication, recombination and repair	Uve	UV damage repair endonuclease	COG04294	pfam03851		TIGR00629		1									
arCOG00115	L	Replication, recombination and repair	-	DNA modification methylase	COG00863	pfam01555											1	1	1
arCOG00792	L	Replication, recombination and repair	-	Superfamily I DNA/RNA helicase	COG01112	pfam13086,pfam13087		TIGR00376									1	1	1
arCOG00919	L	Replication, recombination and repair	-	Holliday junction resolvase, archaeal type	COG01591	pfam01870	cd00523										1	1	2
arCOG02207	L	Replication, recombination and repair	XthA	Exonuclease III	COG00708	pfam03372	cd09073	TIGR00633									1	1	2
arCOG00048	L	Replication, recombination and repair	RlmL	23S rRNA G2445 N2-methylase RlmL	COG00116	pfam02926,pfam0117	cd11715	TIGR01177									1		
arCOG00426	L	Replication, recombination and repair	RecJ	DHH superfamily phosphohydrolase/exonuclease	COG00608	pfam01368											1		
arCOG00463	L	Replication, recombination and repair	-	Uri superfamily endonuclease	COG01833	pfam01986	cd10441										1		
arCOG00928	L	Replication, recombination and repair	-	Endonuclease V homolog	COG01628	pfam01949											1		
arCOG01074	L	Replication, recombination and repair	-	NUDIX family hydrolase	COG00494	pfam00293	cd03427	TIGR02705									1		
arCOG04206	L	Replication, recombination and repair	MUS81	ERCC4-type nuclease	COG01948	pfam02732,pfam12826		TIGR00596									1		
arCOG04295	L	Replication, recombination and repair	Mpg	3-methyladenine DNA glycosylase	COG02094	pfam02245	cd00540	TIGR00567									1		
arCOG04357	L	Replication, recombination and repair	ENDO3c	Thermostable 8-oxoguanine DNA glycosylase	COG01059	pfam00730	cd00056										1		
arCOG04926	L	Replication, recombination and repair	PolB	DNA polymerase elongation subunit (family B)	COG00417												1		
arCOG05336	L	Replication, recombination and repair	HimA	Bacterial nucleoid DNA-binding protein	COG00776	pfam00216	cd13831	TIGR00987									1		
arCOG06233	L	Replication, recombination and repair	-	Toprim domain and Zn-finger domain		pfam01751,pfam0113	cd03362,cd001	TIGR01057									1		
arCOG06670	L	Replication, recombination and repair	-	Nuclease-related protein, NERD family		pfam08378											1		
arCOG14866	L	Replication, recombination and repair	-	HhH domain containing DNA-binding domain		pfam12836											1		
arCOG00280	L	Replication, recombination and repair	-	HerA helicase	COG00433	pfam01935											2		
arCOG01290	L	Replication, recombination and repair	SplB	DNA repair photolyase	COG01533	pfam04055	cd01335										2		
arCOG00329	L	Replication, recombination and repair	PolB2	DNA polymerase PolB2, inactivated	COG00417	pfam00136	cd05531	TIGR00592										1	1
arCOG00802	L	Replication, recombination and repair	RecB	ATP-dependent exoDNAse (exonuclease V) beta subunit (contains helicase an	COG01074	pfam00580,pfam1336	1	TIGR02785										1	1

arCOG02895	L	Replication, recombination and repair	MutS2	DNA structure-specific ATPase involved in suppression of recombination, Mut	COG01193	pfam00488	cd03243	TIGR01069											1	1
arCOG04748	L	Replication, recombination and repair	UvrB	Excinuclease ABC subunit B, helicase	COG00556	pfam04851,pfam1481	cd00046,cd000	TIGR00631											1	1
arCOG00305	L	Replication, recombination and repair	POL4	DNA polymerase IV (family X)	COG01796														1	3
arCOG04694	L	Replication, recombination and repair	UvrA	Excinuclease ABC subunit A, ATPase	COG00178		cd03271,cd032	TIGR00630											1	7
arCOG00462	L	Replication, recombination and repair	MutY	A/G-specific DNA glycosylase	COG01194	pfam00730	cd00056	TIGR01084											1	
arCOG00873	L	Replication, recombination and repair	-	Excinuclease-associated helix-hairpin-helix domain	COG00322	pfam01541,pfam0215	cd10434,cd065	TIGR00194											1	
arCOG03646	L	Replication, recombination and repair	XseB	Exonuclease VII small subunit	COG01722	pfam02609		TIGR01280											1	
arCOG04513	L	Replication, recombination and repair	XseA	Exonuclease VII, large subunit	COG01570	pfam13742,pfam0260	cd04489	TIGR00237											1	
arCOG04754	L	Replication, recombination and repair	Lig	NAD-dependent DNA ligase	COG00272	pfam01653,pfam0312	cd00114,cd000	TIGR00575											1	
arCOG07300	L	Replication, recombination and repair	-	Uncharacterized protein associated with inactivated PolB-like polymerase															1	
arCOG01082	L	Replication, recombination and repair	-	NUDIX family hydrolase	COG00494	pfam00293	cd02885													1
arCOG01305	L	Replication, recombination and repair	-	Topoisomerase DNA binding C4 zinc finger fused to uncharacterized N-termin	COG01637															1
arCOG02896	L	Replication, recombination and repair	MutS	Mismatch repair ATPase (MutS family)	COG00249	pfam01624,pfam0518	cd03284	TIGR01070												1
arCOG02942	L	Replication, recombination and repair	RnhA	Ribonuclease HI	COG00328	pfam13456	cd09279													1
arCOG04753	L	Replication, recombination and repair	UvrC	Excinuclease ABC subunit C	COG00322	pfam01541,pfam0215	cd10434,cd098	TIGR00194												4
arCOG04244	K	Transcription	RPB10	DNA-directed RNA polymerase, subunit N (RpoN/RPB10)	COG01644	pfam01194			1	1		1		1	1	1	1	1		
arCOG04280	K	Transcription	NagC	Transcriptional regulator/sugar kinase	COG01940	pfam00480	cd00012	TIGR00744	1	1		1		1	1	1	1			
arCOG01863	K	Transcription	-	Predicted transcription factor, homolog of eukaryotic MBF1	COG01813	pfam01381	cd00093	TIGR00270	1	1		1			1	1	1	1	1	2
arCOG04061	K	Transcription	EGD2	Transcription factor homologous to NACalpha-BTF3	COG01308		cd14359	TIGR00264	1	2	1	1	1	1	1	1	1			
arCOG01753	K	Transcription	Ssh10b	Archaeal DNA-binding protein	COG01581	pfam01918		TIGR00285	1		1	1		1	1	1	1			2
arCOG01762	K	Transcription	RpoB/Rpo2	DNA-directed RNA polymerase subunit B	COG00085	pfam04563,pfam0456	cd00653	TIGR03670		1	1	1		1	1	1	1	1	1	1
arCOG04257	K	Transcription	RpoC/Rpo3	DNA-directed RNA polymerase subunit A'	COG00086	pfam04997,pfam0062	cd02582	TIGR02390		1	1	1		1	1	1	1	1	1	3
arCOG02613	K	Transcription	-	Chromosome segregation and condensation protein B	COG01386	pfam04079		TIGR00281		1				1	1	1	1	1		
arCOG03112	K	Transcription	-	Fic family protein	COG03177	pfam02661		TIGR02613		1				1	1	1				
arCOG04256	K	Transcription	RpoC/Rpo11	DNA-directed RNA polymerase subunit A''	COG00086	pfam04998	cd06528	TIGR02389		2	1	1		1	1	1	1	1	1	2
arCOG01128	K	Transcription	TenA	Transcriptional activator TenA	COG00819	pfam03070		TIGR04306							1	1				
arCOG00579	K	Transcription	RPB9	DNA-directed RNA polymerase, subunit M/Transcription elongation factor TFI	COG01594			TIGR01384	1	1	1	1		1	1		1	1	1	2
arCOG04366	K	Transcription	-	ArsR family transcriptional regulator	COG04860	pfam09824			1	1	1	1		1	1		1			
arCOG01684	K	Transcription	-	Transcriptional regulator, ArsR family	COG01777	pfam01022	cd00090,cd00090		1	1	1		1	2	1		1	1	1	4
arCOG01361	K	Transcription	ELP3	Histone acetyltransferase	COG01243	pfam04055		TIGR01211	1	1					1		1			
arCOG01016	K	Transcription	Rpo4	DNA-directed RNA polymerase, subunit Rpo4/RpoF	COG01460	pfam03874			1	2	1	1	1	1	1		1	1	1	1
arCOG01920	K	Transcription	NusG	Transcription antiterminator NusG	COG00250	pfam03439	cd09887,cd060	TIGR00405	1	2	1	1		1	1		1	1	1	2
arCOG00675	K	Transcription	RPB7	DNA-directed RNA polymerase, subunit E'/Rpb7	COG01095	pfam03876,pfam0057	cd04331,cd044	TIGR00448	1						1		1	1		
arCOG04077	K	Transcription	Spt4	Transcription elongation factor Spt4/RpoE2, zinc finger protein	COG02093	pfam06093			1						1		1			
arCOG04258	K	Transcription	RPB5	DNA-directed RNA polymerase, subunit H, RpoH/RPB5	COG02012	pfam01191				1	1	1		1	1		1	1	1	1
arCOG01268	K	Transcription	Rpo6/RpoZ	DNA-directed RNA polymerase subunit K/omega	COG01758	pfam01192				1		1	1	1	1		1	1	1	1
arCOG02611	K	Transcription	-	Predicted transcriptional regulator, containd two HTH domains	COG03398	pfam13412,pfam1341	cd00090,cd00090		1	1		1			2	2	1	3	6	
arCOG07561	K	Transcription	-	Predicted membrane-associated trancriptional regulator	COG02512				2	1		1			2	2	2	7	7	
arCOG01580	K	Transcription	-	DNA-binding transcriptional regulator, Lrp family	COG01522	pfam13412,pfam0103	cd00090		1	1	1	1			2		5	2	6	
arCOG05161	K	Transcription	WecD	Acetyltransferase (GNAT) family	COG00454				1	1		1		1	2					
arCOG04111	K	Transcription	RPB11	DNA-directed RNA polymerase, subunit L	COG01761	pfam13656	cd06927		1	2	1	1		1	2		1	1		
arCOG01764	K	Transcription	SPT15	TATA-box binding protein (TBP), component of TFIID and TFIIIB	COG02101	pfam00352,pfam0035	cd04518		2	1	2	1	1	2	3	1	1	1	1	
arCOG01981	K	Transcription	SUA7	Transcription initiation factor TFIIB, Brf1 subunit/Transcription initiation facto	COG01405	pfam08271,pfam0038	cd00043		3	4	3	2	1	1	3	1	2	2	6	
arCOG02099	K	Transcription	TroR	Mn-dependent transcriptional regulator (DtxR family)	COG01321	pfam01325,pfam02742	pfam04023		2	2	1	1		1	3			2	3	
arCOG04060	K	Transcription	-	Predicted transcriptional regulator	COG01709	pfam01381	cd00093		1	1	1	1	1			1	1			
arCOG00770	K	Transcription	DinG	Rad3-related DNA helicase	COG01199	pfam06733,pfam13307		TIGR00604	1		1					1	1	1	2	
arCOG04241	K	Transcription	RpoA/Rpo1	DNA-directed RNA polymerase subunit D	COG00202	pfam01193	cd07030		1			1				1	1	1	1	
arCOG05349	K	Transcription	ArsR	Transcriptional regulator containing HTH domain, ArsR family	COG00640	pfam12840	cd00090			1						1	1			
arCOG01680	K	Transcription	ArsR	Transcriptional regulator containing HTH domain, ArsR family	COG00640	pfam01022	cd00090			2						1	1			
arCOG02038	K	Transcription	-	Sugar-specific transcriptional regulator TrmB	COG01378											1	1	1		
arCOG02037	K	Transcription	-	Sugar-specific transcriptional regulator TrmB	COG01378	pfam01978			1	1	1	1						1	2	
arCOG04686	K	Transcription	VacB	Exoribonuclease R	COG00557	pfam00773		TIGR02063	1		1	1	1					1		
arCOG04341	K	Transcription	RPC12	DNA-directed RNA polymerase, subunit RPC12/RpoP (contains C4-type Zn-fin	COG01996				1											
arCOG03182	K	Transcription	MarR	Transcriptional regulator, MarR family	COG01846	pfam01047	cd00090			1							1		2	
arCOG01057	K	Transcription	HxlR	DNA-binding transcriptional regulator, HxlR family	COG01733	pfam01638	cd00090			1							2	5	3	
arCOG05671	K	Transcription	CsgD	Transcriptional regulator LuxR family	COG02771					1										
arCOG00839	K	Transcription	WecD	Acetyltransferase (GNAT) family	COG00454	pfam13508	cd04301	TIGR01575				1								1
arCOG00608	K	Transcription	-	Predicted transcriptional regulator with C-terminal CBS domains	COG03620	pfam01381,pfam0057	cd00093,cd046	TIGR03070,TIGR01137									1	1		
arCOG04248	K	Transcription	SIR2	NAD-dependent protein deacetylase, SIR2 family	COG00846	pfam02146	cd01413										1	1		
arCOG04377	K	Transcription	-	Transcriptional regulator MarR family, contains HTH domain	COG04738												1	1		
arCOG00002	K	Transcription	-	Transcriptional regulator, PadR family	COG01695	pfam03551		TIGR03433									1			
arCOG00492	K	Transcription	ARO8	Transcriptional regulators containing a DNA-binding HTH domain and an amin	COG01167	pfam00155	cd00609	TIGR03947									1			
arCOG00610	K	Transcription	-	Predicted transcriptional regulator, contains C-terminal CBS domains	COG02524	pfam03444,pfam0047	cd04588	TIGR00331,TIGR01302									1			
arCOG00732	K	Transcription	MarR	Transcriptional regulator, MarR family	COG01846	pfam13601	cd00090										1			
arCOG00743	K	Transcription	-	Predicted transcriptional regulator													1			
arCOG00845	K	Transcription	WecD	Acetyltransferase (GNAT) family	COG00454	pfam00583	cd04301	TIGR01575									1			
arCOG01055	K	Transcription	-	Transcriptional regulator, contains HTH domain	COG03432	pfam14947											1			
arCOG01586	K	Transcription	-	DNA-binding transcriptional regulator, Lrp family	COG01522	pfam13404,pfam1340	cd00090,cd00090										1			
arCOG01683	K	Transcription	ArsR	Transcriptional regulator containing HTH domain, ArsR family	COG00640	pfam01978	cd00090										1			
arCOG01685	K	Transcription	ArsR	Transcriptional regulator containing HTH domain, ArsR family	COG00640	pfam12840	cd00090										1			
arCOG01760	K	Transcription	NusA	Transcription elongation factor	COG00195		cd02134,cd021	TIGR01952									1			
arCOG01804	K	Transcription	-	Transcriptional regulator, contains HTH domain	COG03888	pfam13412	cd00090,cd13553										1			
arCOG02808	K	Transcription	-	Transcriptional regulator, contains HTH domain	COG04742												1			
arCOG03698	K	Transcription	-	Transciptional regulator, contains HTH domain	COG04344	pfam10007											1			
arCOG04081	K	Transcription	TIP49	DNA helicase TIP49, TBP-interacting protein	COG01224	pfam06068	cd00009	TIGR01241,TIGR02397									1			
arCOG04152	K	Transcription	-	Predicted transcriptional regulator	COG01395	pfam01381	cd00093										1			
arCOG04399	K	Transcription	-	Predicted transcriptional regulator	COG01497												1			
arCOG04479	K	Transcription	-	Transcriptional regulator containing an HTH domain fused to a Zn-ribbon	COG03357												1			
arCOG04939	K	Transcription	-	Transcriptional regulator, contains wHTH domain	SC.00517												1			
arCOG07764	K	Transcription	-	Predicted transcriptional regulator	COG01318												1			
arCOG13514	K	Transcription	-	Transcriptional regulator, contains N-terminal RHH domain													1			

arCOG01117	K	Transcription	-	Lrp/AsnC family C-terminal domain	COG01522	pfam01037												2	3	6
arCOG00826	K	Transcription	WecD	Acetyltransferase (GNAT) family	COG00454	pfam00583	cd04301	TIGR01575										2		1
arCOG00001	K	Transcription	-	Transcriptional regulator, PadR family	COG01695	pfam03551												2		
arCOG00998	K	Transcription	LSM1	Small nuclear ribonucleoprotein (snRNP) homolog	COG01958	pfam01423	cd01726											2		
arCOG01345	K	Transcription	-	Predicted transcriptional regulator	COG03388													2		
arCOG02100	K	Transcription	TroR	Mn-dependent transcriptional regulator (DtxR family)	COG01321	pfam01325,pfam02742												2		
arCOG02394	K	Transcription	CheF	Component of chemotaxis system associated with archaellum, contains CheF-	COG02469													2		
arCOG02795	K	Transcription	GbsR	DNA-binding transcriptional regulator GbsR, MarR family	COG01510													2		
arCOG03067	K	Transcription	ArsR	Transcriptional regulator containing HTH domain, ArsR family	COG00640	pfam12840												3		
arCOG02983	K	Transcription	CspC	Cold shock protein, CspA family	COG01278	pfam00313	cd04458	TIGR02381											1	2
arCOG01875	K	Transcription	-	ParB-like nuclease domain	COG01475	pfam02195													1	3
arCOG02644	K	Transcription	AcrR	Transcriptional regulator, TetR/AcrR family	COG01309	pfam00440		TIGR03613											1	
arCOG05152	K	Transcription	-	Transcriptional regulator, contains HTH domain															1	
arCOG03748	K	Transcription	MarR	Transcriptional regulator, MarR family	COG01846														2	1
arCOG02271	K	Transcription	-	Transcriptional regulator, contains HTH domain	COG03413	pfam04967													2	4
arCOG02274	K	Transcription	-	Transcriptional regulator, contains HTH domain	COG03413	pfam04967													3	
arCOG02280	K	Transcription	-	Transcriptional regulator, contains HTH domain	COG03413	pfam04967														1
arCOG04818	K	Transcription	HepA	Superfamily II DNA/RNA helicase, SNF2 family	COG00553	pfam13091,pfam0765	cd09178,cd000	TIGR01587												1
arCOG00398	J	Translation, ribosomal structure and biogenesis	-	ASCH domain, predicted RNA-binding domain	COG02411	pfam04266	cd06552											1		
arCOG00423	J	Translation, ribosomal structure and biogenesis	-	Oligoribonuclease NrnB or cAMP/cGMP phosphodiesterase, DHH superfamily	COG02404	pfam01368,pfam02272												1		
arCOG00910	J	Translation, ribosomal structure and biogenesis	-	Predicted RNA methylase	COG02263	pfam13659						1	1	1	1	1		1	1	1
arCOG00953	J	Translation, ribosomal structure and biogenesis	-	Homolog of Wybutosine (yW) biosynthesis enzyme, Fe-S oxidoreductase	COG00731	pfam04055												1		
arCOG00990	J	Translation, ribosomal structure and biogenesis	-	Queuine tRNA-ribosyltransferase, contain PUA domain	COG01549	pfam14810,pfam01472		TIGR00432	1	1	1	1			1			1	1	2
arCOG00991	J	Translation, ribosomal structure and biogenesis	-	tRNA modification protein, contains pre-PUA and PUA domains	COG01370	pfam14810,pfam01472		TIGR00432	1	1					1	1		1	1	2
arCOG01042	J	Translation, ribosomal structure and biogenesis	-	Exosome subunit, RNA binding protein with dsRBD fold	COG01325	pfam01877												1	1	1
arCOG01346	J	Translation, ribosomal structure and biogenesis	-	RNA-binding protein containing KH domain, possibly ribosomal protein	COG01534	pfam01985												1	1	2
arCOG01695	J	Translation, ribosomal structure and biogenesis	-	Predicted component of the ribosome quality control (RQC) complex, YloA/Ta	COG01293	pfam05833,pfam05670			1						1	1		1	2	2
arCOG01861	J	Translation, ribosomal structure and biogenesis	-	3'-5' exoribonuclease YhaM, can participate in 23S rRNA maturation, HD supe	COG03481	pfam01966		TIGR00277										2		
arCOG04130	J	Translation, ribosomal structure and biogenesis	-	Predicted RNA-binding protein	COG01491	pfam04919			1	2	1	1	1	1	1			1	1	1
arCOG04318	J	Translation, ribosomal structure and biogenesis	-	Predicted RNA-binding protein of the translin family	COG02178	pfam01997	cd14820											1		
arCOG01255	J	Translation, ribosomal structure and biogenesis	AlaS	Alanyl-tRNA synthetase	COG00013	pfam01411,pfam0797	cd00673,cd077	TIGR03683	1	2	1	1			1	1		2	1	1
arCOG01254	J	Translation, ribosomal structure and biogenesis	AlaX	Ser-tRNA(Ala) deacylase AlaX (editing enzyme)	COG02872	pfam01411,pfam07973		TIGR03683			1	1			1			1	1	2
arCOG01256	J	Translation, ribosomal structure and biogenesis	AlaX	Ser-tRNA(Ala) deacylase AlaX (editing enzyme)	COG02872			TIGR03683										1		
arCOG00487	J	Translation, ribosomal structure and biogenesis	ArgS	Arginyl-tRNA synthetase	COG00018	pfam00750,pfam0574	cd00671,cd079	TIGR00456	1		1	1	1	1	1			1	1	1
arCOG00406	J	Translation, ribosomal structure and biogenesis	AsnS	Aspartyl/asparaginyl-tRNA synthetase	COG00017	pfam01336,pfam0015	cd04316,cd007	TIGR00458			1		1				1	1	1	1
arCOG00407	J	Translation, ribosomal structure and biogenesis	AsnS	Aspartyl/asparaginyl-tRNA synthetase	COG00017	pfam01336,pfam0015	cd04319,cd007	TIGR00457	1	2	1	1	1	1	1			1	1	2
arCOG01185	J	Translation, ribosomal structure and biogenesis	Bud32	tRNA A-37 threonylcarbamoyl transferase component Bud32	COG03642	pfam00069	cd05151	TIGR03724	2	2	1	1		1	2			1	2	4
arCOG04249	J	Translation, ribosomal structure and biogenesis	CCA1	tRNA nucleotidyltransferase (CCA-adding enzyme)	COG01746	pfam01909,pfam0924	cd05400	TIGR03671	1	1	1		1	2	1	1		1		
arCOG02197	J	Translation, ribosomal structure and biogenesis	Cgi121	Subunit of KEOPS complex (Cgi121BUD32KAE1)	COG01617	pfam08617			1	1		1			1			1		
arCOG00676	J	Translation, ribosomal structure and biogenesis	Csl4	Exosome complex RNA-binding protein Csl4, contains S1 and Zn-ribbon doma	COG01096	pfam14382,pfam0057	cd05692					1	1	1	1	1		1	1	1
arCOG00486	J	Translation, ribosomal structure and biogenesis	CysS	Cysteinyl-tRNA synthetase	COG00215	pfam01406,pfam0919	cd00672,cd079	TIGR00435	1	1		2		2	3	2		1	1	2
arCOG04112	J	Translation, ribosomal structure and biogenesis	DPH2	Diphthamide synthase subunit DPH2	COG01736	pfam01866		TIGR00322	1	1	1	1	1	1	2			1	1	2
arCOG04161	J	Translation, ribosomal structure and biogenesis	DPH5	Diphthamide biosynthesis methyltransferase	COG01798	pfam00590	cd11647	TIGR00522	1		1		1			1		1	1	2
arCOG00035	J	Translation, ribosomal structure and biogenesis	Dph6	Diphthamide synthase (EF-2-diphthine--ammonia ligase)	COG02102	pfam01902	cd01994	TIGR03679	1	1		1			1			1	1	1
arCOG00358	J	Translation, ribosomal structure and biogenesis	DRG	Ribosome-interacting GTPase 1	COG01163	pfam02824	cd01896,cd01666		1	2	1	1	1		2			1	1	1
arCOG00604	J	Translation, ribosomal structure and biogenesis	DusA	tRNA-dihydrouridine synthase	COG00042	pfam01207	cd02801	TIGR00737											1	3
arCOG04332	J	Translation, ribosomal structure and biogenesis	EbsC	Cys-tRNA(Pro)/Cys-tRNA(Cys) deacylase, ybaK family	COG02606	pfam04073	cd04333	TIGR00011												2
arCOG01988	J	Translation, ribosomal structure and biogenesis	EFB1	Translation elongation factor EF-1beta	COG02092	pfam00736	cd00292	TIGR00489	1	1	1	1		1	1			1	1	
arCOG04277	J	Translation, ribosomal structure and biogenesis	Efp	Translation elongation factor P (EF-P)/translation initiation factor 5A (eIF-5A)	COG00231	pfam08207,pfam0128	cd04467	TIGR00037	1		1			1	1	1		1	1	
arCOG04122	J	Translation, ribosomal structure and biogenesis	Emg1	rRNA pseudouridine-1189 N-methylase Emg1, Nep1/Mra1 family	COG01756	pfam03587												1		
arCOG01742	J	Translation, ribosomal structure and biogenesis	eRF1	Peptide chain release factor eRF1	COG01503	pfam03463,pfam03464	pfam03465	TIGR03676					1			1		1	1	1
arCOG04312	J	Translation, ribosomal structure and biogenesis	Fcf1	rRNA-processing protein FCF1, contains PIN domain	COG01412		cd09879		1		1	1		1	1	1		1	1	2
arCOG00079	J	Translation, ribosomal structure and biogenesis	FtsJ	23S rRNA U2552 (ribose-2'-O)-methylase RlmE/FtsJ	COG00293	pfam01728	cd02440	TIGR00438	1	1		1			1			1	1	1
arCOG01559	J	Translation, ribosomal structure and biogenesis	FusA	Translation elongation factor G, EF-G (GTPase)	COG00480	pfam00009,pfam0314	cd01885,cd037	TIGR00490	1	1	1	1		1	1			1	1	2
arCOG02466	J	Translation, ribosomal structure and biogenesis	GAR1	RNA-binding protein involved in rRNA processing	COG03277				1	1	1	1			1			1		
arCOG01719	J	Translation, ribosomal structure and biogenesis	GatE	Archaeal Glu-tRNAGln amidotransferase subunit E (contains GAD domain)	COG02511	pfam02934,pfam0263	7	TIGR00134	1	1	1	1	1		1	1		1	1	3
arCOG01563	J	Translation, ribosomal structure and biogenesis	GCD11	Translation initiation factor 2, gamma subunit eIF-2 gamma, GTPase	COG05257	pfam00009,pfam0917	cd01888,cd036	TIGR03680	1		1	1		1	1	1		1	1	4
arCOG00978	J	Translation, ribosomal structure and biogenesis	GCD14	tRNA(1-methyladenosine) methyltransferase	COG02519	pfam08704	cd02440	TIGR03534	1	1	1		1	2	1	1		1	1	1
arCOG01124	J	Translation, ribosomal structure and biogenesis	GCD2	Translation initiation factor 2B subunit, eIF-2B alpha/beta/delta family	COG01184	pfam01008		TIGR00511										1		
arCOG01640	J	Translation, ribosomal structure and biogenesis	GCD7	Translation initiation factor 2, beta subunit (eIF-2beta)/eIF-5 N-terminal doma	COG01601	pfam01873		TIGR00311								1				
arCOG01616	J	Translation, ribosomal structure and biogenesis	GEK1	D-aminoacyl-tRNA deacylase, involved in ethanol tolerance	COG01650	pfam04414			1	1	1	1			1			1	1	
arCOG04302	J	Translation, ribosomal structure and biogenesis	GlnS	Glutamyl- or glutaminyl-tRNA synthetase	COG00008	pfam00749,pfam0395	cd09287	TIGR00463	2	2	1	1	1	1	1			1	1	2
arCOG00405	J	Translation, ribosomal structure and biogenesis	GRS1	Glycyl-tRNA synthetase (class II)	COG00423	pfam00587,pfam0312	cd00774,cd008	TIGR00389		1	1				1	1		1	1	1
arCOG00357	J	Translation, ribosomal structure and biogenesis	GTP1	Ribosome-binding ATPase YchF, GTP1/OBG family	COG00012	pfam01926,pfam0843	cd01899,cd016	TIGR02729	1						2			1	1	2
arCOG00109	J	Translation, ribosomal structure and biogenesis	HemK	Methylase of polypeptide chain release factors	COG02890	pfam13659	cd02440	TIGR00537	1	2					1	1		1		
arCOG00404	J	Translation, ribosomal structure and biogenesis	HisS	Histidyl-tRNA synthetase	COG00124	pfam13393,pfam0312	cd00773,cd008	TIGR00442	1	1					1			1	1	2
arCOG00807	J	Translation, ribosomal structure and biogenesis	IleS	Isoleucyl-tRNA synthetase	COG00060	pfam00133,pfam0826	cd00818,cd008	TIGR00392		1				1				1	1	4
arCOG01179	J	Translation, ribosomal structure and biogenesis	InfA	Translation initiation factor 1 (IF-1)	COG00361	pfam01176	cd05793	TIGR00523	1	1		1						1	1	2
arCOG01560	J	Translation, ribosomal structure and biogenesis	InfB	Translation initiation factor 2 (IF-2; GTPase)	COG00532	pfam00009,pfam0314	cd01887,cd037	TIGR00491	1	2	1	1		1	1	1		1	1	
arCOG01183	J	Translation, ribosomal structure and biogenesis	Kae1p/TsaD	Subunit of KEOPS complex, tRNA A37 threonylcarbamoyltransferase, contains	COG00533	pfam00814		TIGR03722			1	1		1						2
arCOG04063	J	Translation, ribosomal structure and biogenesis	KptA	RNA:NAD 2'-phosphotransferase	COG01859	pfam01885			1	2	1	1		1	1	1		1	1	2
arCOG04150	J	Translation, ribosomal structure and biogenesis	Krr1	rRNA processing protein Krr1/Pno1, contains KH domain	COG01094			TIGR03665	1	1		1						1	1	2
arCOG00809	J	Translation, ribosomal structure and biogenesis	LeuS	Leucyl-tRNA synthetase	COG00495	pfam00133,pfam0826	cd00812,cd008	TIGR00395	1	1								1		
arCOG01736	J	Translation, ribosomal structure and biogenesis	LigT	2'-5' RNA ligase	COG01514	pfam13563		TIGR02258								1		1		
arCOG00485	J	Translation, ribosomal structure and biogenesis	LysS	Lysyl-tRNA synthetase, class I	COG01384	pfam01921	cd00674	TIGR00467	1		1				1	1			1	1
arCOG00408	J	Translation, ribosomal structure and biogenesis	LysU	Lysyl-tRNA synthetase (class II)	COG01190	pfam01336,pfam0015	cd04322,cd007	TIGR00499										1		
arCOG01001	J	Translation, ribosomal structure and biogenesis	Map	Methionine aminopeptidase	COG00024	pfam00557	cd01088	TIGR00501	1	2	1	1	1	1	2			1	1	2
arCOG00810	J	Translation, ribosomal structure and biogenesis	MetG	Methionyl-tRNA synthetase	COG00143	pfam09334,pfam0826	cd00814,cd079	TIGR00398	1		1			1		1		1	1	2
arCOG01358	J	Translation, ribosomal structure and biogenesis	MiaB	2-methylthioadenine synthetase	COG00621	pfam00919,pfam0405	cd01335	TIGR00089	1									1		
arCOG02286	J	Translation, ribosomal structure and biogenesis	Ncl1	Ribosome biogenesis protein, NOL1/NOP2/fmu family	COG03270	pfam13636			1	1	1	1			1			1		
arCOG04149	J	Translation, ribosomal structure and biogenesis	NMD3	NMD protein affecting ribosome stability and mRNA decay	COG01499	pfam04981			1	1				1	1	2		1	1	1

arCOG00721	J	Translation, ribosomal structure and biogenesis	Nob1	Endonuclease Nob1, consists of a PIN domain and a Zn-ribbon module	COG01439		cd09876		1						1		1	1	1
arCOG00078	J	Translation, ribosomal structure and biogenesis	NOP1	Fibrillarin-like rRNA methylase	COG01889	pfam01269			1	1	1						1	1	1
arCOG00906	J	Translation, ribosomal structure and biogenesis	Nop10	rRNA maturation protein Nop10, contains Zn-ribbon domain	COG02260	pfam04135			1	1		1	1	1	1		1	1	2
arCOG04414	J	Translation, ribosomal structure and biogenesis	Pcc1	Subunit of KEOPS complex (Cgi121BUD32KAE1)	COG02892				1		1		1		1		1		
arCOG01741	J	Translation, ribosomal structure and biogenesis	PelA	Release factor eRF1	COG01537	pfam03463,pfam03464,pfam03465		TIGR00111	1	2	1	1	1	1	1	1	1	1	2
arCOG00410	J	Translation, ribosomal structure and biogenesis	PheS	Phenylalanyl-tRNA synthetase alpha subunit	COG00016	pfam01409	cd00496	TIGR00468	1	2	1				1	1	1	1	1
arCOG00412	J	Translation, ribosomal structure and biogenesis	PheT	Phenylalanyl-tRNA synthetase beta subunit	COG00072	pfam03483,pfam03484	cd00769	TIGR00471	1	2	1		1	1	1		1	1	2
arCOG00784	J	Translation, ribosomal structure and biogenesis	POP4	RNase P/RNase MRP subunit p29	COG01588	pfam01868			1	2	1					1	1		2
arCOG01365	J	Translation, ribosomal structure and biogenesis	POP5	RNase P/RNase MRP subunit POP5	COG01369	pfam01900											1		
arCOG00402	J	Translation, ribosomal structure and biogenesis	ProS	Prolyl-tRNA synthetase	COG00442	pfam00587,pfam03129	cd00778,cd00	TIGR00408	1	1	1	1	1	1	1		1	1	3
arCOG04228	J	Translation, ribosomal structure and biogenesis	Pth2	Peptidyl-tRNA hydrolase	COG01990	pfam01981	cd02430	TIGR00283	1	1	1	1		1	1		1	2	2
arCOG01015	J	Translation, ribosomal structure and biogenesis	Pus10	tRNA U54 and U55 pseudouridine synthase Pus10	COG01258			TIGR01213	1	3	1	1	1	1	1		1	1	1
arCOG13558	J	Translation, ribosomal structure and biogenesis	QueA	S-adenosylmethionine:tRNA ribosyltransferase-isomerase		pfam02547		TIGR00113									1		
arCOG00039	J	Translation, ribosomal structure and biogenesis	QueC	7-cyano-7-deazaguanine synthase (queuosine biosynthesis)	COG00603	pfam06508	cd01995	TIGR00364		1						1	1	1	1
arCOG04210	J	Translation, ribosomal structure and biogenesis	QueFC	NADPH-dependent 7-cyano-7-deazaguanine reductase QueF, C-terminal dom	COG00780	pfam14489		TIGR03139									1		
arCOG04125	J	Translation, ribosomal structure and biogenesis	RCL1	RNA 3'-terminal phosphate cyclase	COG00430	pfam01137	cd00874	TIGR03399									1		
arCOG00842	J	Translation, ribosomal structure and biogenesis	RimL	Acetyltransferase, RimL family	COG01670	pfam13302			1	1						1	1		
arCOG00187	J	Translation, ribosomal structure and biogenesis	Rli1	Translation initiation factor RLI1, contains Fe-S and AAA+ ATPase domains	COG01245	pfam00005,pfam0000	cd03237,cd032	TIGR03269	1	2	1	1		1	1	1	1	1	6
arCOG05111	J	Translation, ribosomal structure and biogenesis	RlmH	23S rRNA pseudoU1915 N3-methylase RlmH	COG01576	pfam02590		TIGR00246									1	1	
arCOG06660	J	Translation, ribosomal structure and biogenesis	RluA	Pseudouridylate synthase, 23S RNA-specific	COG00564	pfam00849	cd02869	TIGR00005										1	2
arCOG00547	J	Translation, ribosomal structure and biogenesis	RnjA	mRNA degradation ribonuclease J1/J2	COG00595	pfam12706		TIGR03675,TIGR02651									1		
arCOG00501	J	Translation, ribosomal structure and biogenesis	Rnz	Ribonuclease Z, beta-lactamase superfamily hydrolase	COG01234	pfam12706		TIGR02651	1	1		1					1	1	2
arCOG01575	J	Translation, ribosomal structure and biogenesis	Rph	Ribonuclease PH	COG00689	pfam01138,pfam0372	cd11366	TIGR02065	1	1	1		1		1		2	2	4
arCOG04209	J	Translation, ribosomal structure and biogenesis	RPL15A	Ribosomal protein L15E	COG01632	pfam00827					1	1			1	1	1	1	3
arCOG00780	J	Translation, ribosomal structure and biogenesis	RPL18A	Ribosomal protein L18E	COG01727	pfam00828		TIGR01071	1	1		1		1	1	1	1	1	
arCOG04089	J	Translation, ribosomal structure and biogenesis	RPL19A	Ribosomal protein L19E	COG02147	pfam01280	cd01418		1	2	1			1	1	1	1	1	1
arCOG04175	J	Translation, ribosomal structure and biogenesis	RPL20A	Ribosomal protein L20A (L18A)	COG02157	pfam01775											1	1	
arCOG04129	J	Translation, ribosomal structure and biogenesis	RPL21A	Ribosomal protein L21E	COG02139	pfam01157			1	2	1	1	1	1	1		1	1	1
arCOG01950	J	Translation, ribosomal structure and biogenesis	RPL24A	Ribosomal protein L24E	COG02075	pfam01246	cd00472		1	1				1	1	1	1		
arCOG01752	J	Translation, ribosomal structure and biogenesis	RPL30	Ribosomal protein L30E	COG01911	pfam01248				1	1	1		1	1	1	1	1	2
arCOG04473	J	Translation, ribosomal structure and biogenesis	RPL31A	Ribosomal protein L31E	COG02097	pfam01198	cd00463		1	2	1	1	1	1	1	1	1	1	1
arCOG00781	J	Translation, ribosomal structure and biogenesis	RPL32	Ribosomal protein L32E	COG01717	pfam01655	cd00513		1	2	1			1	1	1	1	1	1
arCOG04177	J	Translation, ribosomal structure and biogenesis	RPL39	Ribosomal protein L39E	COG02167	pfam00832											1		1
arCOG04049	J	Translation, ribosomal structure and biogenesis	RPL40A	Ribosomal protein L40E	COG01552							1							1
arCOG04109	J	Translation, ribosomal structure and biogenesis	RPL42A	Ribosomal protein L44E	COG01631	pfam00935			1	1		1	1	1	1		1		2
arCOG04208	J	Translation, ribosomal structure and biogenesis	RPL43A	Ribosomal protein L37AE/L43A	COG01997	pfam01780		TIGR00280	1	1	1				1		1		2
arCOG01751	J	Translation, ribosomal structure and biogenesis	Rpl7Ae	Ribosomal protein L7AE	COG01358	pfam01248		TIGR03677	1	1				1	1	1	1	1	1
arCOG04289	J	Translation, ribosomal structure and biogenesis	RplA	Ribosomal protein L1	COG00081	pfam00687	cd00403	TIGR01169	1	1	1				1	1	1	1	2
arCOG04067	J	Translation, ribosomal structure and biogenesis	RplB	Ribosomal protein L2	COG00090	pfam00181,pfam03947		TIGR01171	1	1	1			1		1	1	1	2
arCOG04070	J	Translation, ribosomal structure and biogenesis	RplC	Ribosomal protein L3	COG00087	pfam00297		TIGR03626	1	1	1					1	1	1	2
arCOG04071	J	Translation, ribosomal structure and biogenesis	RplD	Ribosomal protein L4	COG00088	pfam00573		TIGR03672	1	1	1					1	1	1	2
arCOG04092	J	Translation, ribosomal structure and biogenesis	RplE	Ribosomal protein L5	COG00094	pfam00281,pfam00673			1	2	1			1	1	1	1	1	1
arCOG04090	J	Translation, ribosomal structure and biogenesis	RplF	Ribosomal protein L6P	COG00097	pfam00347,pfam00347		TIGR03653	1	2	1			1	1	1	1	1	1
arCOG04288	J	Translation, ribosomal structure and biogenesis	RplJ	Ribosomal protein L10	COG00244	pfam00466	cd05795		1	1	1				1	1	1	1	2
arCOG04372	J	Translation, ribosomal structure and biogenesis	RplK	Ribosomal protein L11	COG00080	pfam03946,pfam0029	cd00349	TIGR01632	1	2	1	1		1	1		1	1	2
arCOG04242	J	Translation, ribosomal structure and biogenesis	RplM	Ribosomal protein L13	COG00102	pfam00572	cd00392	TIGR01077	1	1		1		1	1	1	1	1	
arCOG04095	J	Translation, ribosomal structure and biogenesis	RplN	Ribosomal protein L14	COG00093	pfam00238		TIGR03673	1	2	1			1		1	1	1	2
arCOG00779	J	Translation, ribosomal structure and biogenesis	RplO	Ribosomal protein L15	COG00200	pfam00828			1	2	1			1	1	1	1	1	1
arCOG04113	J	Translation, ribosomal structure and biogenesis	RplP	Ribosomal protein L10AE/L16	COG00197	pfam00252	cd01433	TIGR00279	1	1	1	1			1		1	1	1
arCOG04088	J	Translation, ribosomal structure and biogenesis	RplR	Ribosomal protein L18	COG00256	pfam00861,pfam1420	cd00432		1	2	1			1	1	1	1	1	1
arCOG04098	J	Translation, ribosomal structure and biogenesis	RplV	Ribosomal protein L22	COG00091	pfam00237	cd00336	TIGR01038	1	1	1			1		1	1	1	3
arCOG04072	J	Translation, ribosomal structure and biogenesis	RplW	Ribosomal protein L23	COG00089	pfam00276		TIGR03636	1	1	1			1		1	1	1	2
arCOG04094	J	Translation, ribosomal structure and biogenesis	RplX	Ribosomal protein L24	COG00198		cd06089	TIGR01080	1	2	1			1		1	1		2
arCOG00785	J	Translation, ribosomal structure and biogenesis	RpmC	Ribosomal protein L29	COG00255	pfam00831	cd00427	TIGR00012	1	2	1			1		1	1		
arCOG04086	J	Translation, ribosomal structure and biogenesis	RpmD	Ribosomal protein L30	COG01841		cd01657	TIGR01309	1	2	1			1	1	1	1	1	1
arCOG04287	J	Translation, ribosomal structure and biogenesis	RPP1A	Ribosomal protein L12E/L44/L45/RPP1/RPP2	COG02058	pfam00428	cd05832	TIGR03685	1	1	1				1	1	1	1	2
arCOG04345	J	Translation, ribosomal structure and biogenesis	RPR2	RNase P subunit RPR2	COG02023	pfam04032				1		1				1	1		1
arCOG01885	J	Translation, ribosomal structure and biogenesis	RPS17A	Ribosomal protein S17E	COG01383	pfam00833			1	1	1	1	1	1	2	1	1	1	1
arCOG01344	J	Translation, ribosomal structure and biogenesis	RPS19A	Ribosomal protein S19E (S16A)	COG02238	pfam01090			1	2	1	1	1	1	1	1	1	1	1
arCOG04186	J	Translation, ribosomal structure and biogenesis	RPS1A	Ribosomal protein S3AE	COG01890	pfam01015			1						1	1	1	1	3
arCOG04182	J	Translation, ribosomal structure and biogenesis	RPS24A	Ribosomal protein S24E	COG02004	pfam01282			1						1		1	1	
arCOG04108	J	Translation, ribosomal structure and biogenesis	RPS27A	Ribosomal protein S27E	COG02051	pfam01667			1	1		1	1	1			1		1
arCOG04183	J	Translation, ribosomal structure and biogenesis	RPS27AE	Ribosomal protein S27AE	COG01998				1								1		
arCOG04314	J	Translation, ribosomal structure and biogenesis	RPS28A	Ribosomal protein S28E/S33	COG02053	pfam01200	cd04457		1	1				1	1	1	1	1	1
arCOG04093	J	Translation, ribosomal structure and biogenesis	RPS4A	Ribosomal protein S4E	COG01471	pfam08071,pfam0147	cd00165,cd06087		1	2	1			1	1	1	1	1	2
arCOG01946	J	Translation, ribosomal structure and biogenesis	RPS6A	Ribosomal protein S6E/S10	COG02125	pfam01092			1		1	1		1	1	1	1	1	3
arCOG04154	J	Translation, ribosomal structure and biogenesis	RPS8A	Ribosomal protein S8E	COG02007	pfam01201	cd11382	TIGR00307	1	1		1			1	1	1	1	2
arCOG04245	J	Translation, ribosomal structure and biogenesis	RpsB	Ribosomal protein S2	COG00052	pfam00318	cd01425	TIGR01012		1		1	1		1		1	1	2
arCOG04097	J	Translation, ribosomal structure and biogenesis	RpsC	Ribosomal protein S3	COG00092	pfam07650,pfam0018	cd02411	TIGR01008	1	2	1			1		1	1	1	2
arCOG04239	J	Translation, ribosomal structure and biogenesis	RpsD	Ribosomal protein S4 or related protein	COG00522	pfam00163,pfam0147	cd00165	TIGR01018	1			1				1	1	1	1
arCOG04087	J	Translation, ribosomal structure and biogenesis	RpsE	Ribosomal protein S5	COG00098	pfam00333,pfam03719		TIGR01020	1	2	1			2	1	1	1	1	1
arCOG04254	J	Translation, ribosomal structure and biogenesis	RpsG	Ribosomal protein S7	COG00049	pfam00177	cd14867	TIGR01028	1	1	1	1		1	1		1	1	1
arCOG04091	J	Translation, ribosomal structure and biogenesis	RpsH	Ribosomal protein S8	COG00096	pfam00410			1	2	1			1	1	1	1	1	1
arCOG04243	J	Translation, ribosomal structure and biogenesis	RpsI	Ribosomal protein S9	COG00103	pfam00380		TIGR03627	1	1		1		1	1	1	1	1	
arCOG01758	J	Translation, ribosomal structure and biogenesis	RpsJ	Ribosomal protein S10	COG00051	pfam00338		TIGR01046	1	1		1		1	1		1	1	2
arCOG04240	J	Translation, ribosomal structure and biogenesis	RpsK	Ribosomal protein S11	COG00100	pfam00411		TIGR03628	1			1				1	1	1	1
arCOG04255	J	Translation, ribosomal structure and biogenesis	RpsL	Ribosomal protein S12	COG00048	pfam00164	cd03367	TIGR00982	1	1	1	1		1	1		1	1	1
arCOG01722	J	Translation, ribosomal structure and biogenesis	RpsM	Ribosomal protein S13	COG00099	pfam00416		TIGR03629	1			1				1	1	1	1
arCOG00782	J	Translation, ribosomal structure and biogenesis	RpsN	Ribosomal protein S14	COG00199	pfam00253			1	1	1			1			1		
arCOG04185	J	Translation, ribosomal structure and biogenesis	RpsO	Ribosomal protein S15P	COG00184	pfam08069,pfam0031	cd00353		1	1			1	2	1	1	1	1	2
arCOG04096	J	Translation, ribosomal structure and biogenesis	RpsQ	Ribosomal protein S17	COG00186	pfam00366		TIGR03630	1	2	1			1		1	1	1	2

arCOG04099	J	Translation, ribosomal structure and biogenesis	RpsS	Ribosomal protein S19	COG00185	pfam00203		TIGR01025	1	1	1			1		1	1	1	2
arCOG00678	J	Translation, ribosomal structure and biogenesis	Rrp4	Exosome complex RNA-binding protein Rrp4, contains S1 and KH domains	COG01097		cd05789		1	1	1		1		1		1	1	1
arCOG01574	J	Translation, ribosomal structure and biogenesis	Rrp42	Exosome complex RNA-binding protein Rrp42, RNase PH superfamily	COG02123	pfam01138,pfam0372	cd11365	TIGR01966	1	1	1		1		1		1	1	2
arCOG04131	J	Translation, ribosomal structure and biogenesis	RsmA	16S rRNA A1518 and A1519 N6-dimethyltransferase RsmA/KsgA/DIM1	COG00030	pfam00398		TIGR00755	1	1	1				1		1	1	1
arCOG00973	J	Translation, ribosomal structure and biogenesis	RsmB	16S rRNA C967 or C1407 C5-methylase, RsmB/RsmF family	COG00144	pfam01189		TIGR00446	1	1	1	1			1		1	1	3
arCOG00975	J	Translation, ribosomal structure and biogenesis	RsmB	16S rRNA C967 or C1407 C5-methylase, RsmB/RsmF family	COG00144	pfam01189	cd02440	TIGR00446									1		
arCOG01239	J	Translation, ribosomal structure and biogenesis	RsmE	RNA base methyltransferase family enzyme	COG01901	pfam04013			1	2	1	1			1	1	1	1	1
arCOG04246	J	Translation, ribosomal structure and biogenesis	RtcB	RNA 3'-P ligase, RtcB family protein	COG01690	pfam01139		TIGR03073							1		1	1	1
arCOG04187	J	Translation, ribosomal structure and biogenesis	Sdo1	Ribosome maturation protein Sdo1	COG01500	pfam01172,pfam09377		TIGR00291	1	1	1		1		1		1	1	1
arCOG01564	J	Translation, ribosomal structure and biogenesis	SelB	Selenocysteine-specific translation elongation factor or SelB-II domain	COG03276		cd03696	TIGR00475	1	1					1			1	3
arCOG01701	J	Translation, ribosomal structure and biogenesis	SEN2	tRNA splicing endonuclease	COG01676	pfam02778,pfam01974		TIGR00324	1		1				1	1	1	1	5
arCOG00403	J	Translation, ribosomal structure and biogenesis	SerS	Seryl-tRNA synthetase	COG00172	pfam02403,pfam0058	cd00770	TIGR00414	1			1		1	1	1	1	1	1
arCOG01923	J	Translation, ribosomal structure and biogenesis	SIK1	RNA processing factor Prp31, contains Nop domain	COG01498	pfam01798			1	1	1	1		1		1	1	1	2
arCOG01952	J	Translation, ribosomal structure and biogenesis	SUA5	tRNA A37 threonylcarbamoyladenosine synthetase subunit TsaC/SUA5/YrdC	COG00009	pfam01300		TIGR00057	1	2	1	1	1	1	2		1	1	1
arCOG04223	J	Translation, ribosomal structure and biogenesis	SUI1	Translation initiation factor 1 (eIF-1/SUI1)	COG00023	pfam01253	cd11567	TIGR01158	1	2	1			1		1	1	1	2
arCOG04107	J	Translation, ribosomal structure and biogenesis	SUI2	Translation initiation factor 2, alpha subunit (eIF-2alpha)	COG01093	pfam00575,pfam0754	cd04452	TIGR00717	1	1		1	1	1	1		1	1	2
arCOG00056	J	Translation, ribosomal structure and biogenesis	Tan1	tRNA(Ser,Leu) C12 N-acetylase TAN1, contains THUMP domain	COG01818	pfam02926,pfam1441	cd11718,cd061	TIGR00342							1				
arCOG01630	J	Translation, ribosomal structure and biogenesis	TdcF	Translation initiation inhibitor, yjgF family	COG00251	pfam01042	cd00448	TIGR00004	2	2	2			1	3	2		2	4
arCOG01561	J	Translation, ribosomal structure and biogenesis	TEF1	Translation elongation factor EF-1 alpha, GTPase	COG05256	pfam00009,pfam0314	cd01883,cd036	TIGR00483	1	1	1	1		1	1		1	1	5
arCOG00989	J	Translation, ribosomal structure and biogenesis	Tgt	Queuine/archaeosine tRNA-ribosyltransferase	COG00343	pfam01702		TIGR00432									1		
arCOG00038	J	Translation, ribosomal structure and biogenesis	ThiI	tRNA S(4)U 4-thiouridine synthase	COG00301	pfam02926,pfam0256	cd11716,cd017	TIGR00342,TIGR04271			1					1	2	2	2
arCOG00401	J	Translation, ribosomal structure and biogenesis	ThrS	Threonyl-tRNA synthetase	COG00441	pfam00587,pfam0312	cd00771,cd008	TIGR00418	1	1				1	1		1	1	4
arCOG01115	J	Translation, ribosomal structure and biogenesis	TiaS	tRNA(Ile2) 2-agmatinylcytidine synthetase; containing Zn-ribbon domain and	COG01571	pfam08489,pfam0728	cd04482	TIGR03280	1	1		1			1	1	1	1	1
arCOG04176	J	Translation, ribosomal structure and biogenesis	TIF6	Translation initiation factor 6 (eIF-6)	COG01976	pfam01912	cd00527	TIGR00323	1	2	1	1	1	1	1	1	1	1	1
arCOG00042	J	Translation, ribosomal structure and biogenesis	TilS	tRNA(Ile)-lysidine synthase TilS/MesJ	COG00037	pfam01171	cd01993	TIGR02432	1	1		1					2	1	3
arCOG00985	J	Translation, ribosomal structure and biogenesis	Tma20	Predicted RNA-binding protein, contains PUA domain	COG02016	pfam09183,pfam01472		TIGR03684	1	1				1	1	1	1	1	2
arCOG01951	J	Translation, ribosomal structure and biogenesis	TmcA	tRNA(Met) C34 N-acetyltransferase TmcA	COG01444	pfam08351,pfam05127,pfam13718											1		
arCOG01219	J	Translation, ribosomal structure and biogenesis	TRM1	N2,N2-dimethylguanosine tRNA methyltransferase	COG01867	pfam02005		TIGR00308	1	2	1	1		1	2		1	1	1
arCOG00047	J	Translation, ribosomal structure and biogenesis	Trm11	tRNA G10 N-methylase Trm11	COG01041	pfam01170		TIGR01177	1			1				1	1	1	2
arCOG00033	J	Translation, ribosomal structure and biogenesis	Trm5	Wybutosine (yW) biosynthesis enzyme, Trm5 methyltransferase	COG02520	pfam02475			2	1	1	2		1		1	2	1	2
arCOG01018	J	Translation, ribosomal structure and biogenesis	TrmJ	tRNA C32,U32 (ribose-2'-O)-methylase TrmJ or a related methyltransferase	COG00565	pfam00588		TIGR00050	1	2	1	1	1	2	1	1	1	2	2
arCOG01887	J	Translation, ribosomal structure and biogenesis	TrpS	Tryptophanyl-tRNA synthetase	COG00180	pfam00579	cd00806	TIGR00233	1	1	1	1	1		1		1	1	3
arCOG04449	J	Translation, ribosomal structure and biogenesis	TruA	Pseudouridylate synthase	COG00101	pfam01416,pfam0141	cd02866	TIGR00071	1	1	1	1					1	1	2
arCOG00987	J	Translation, ribosomal structure and biogenesis	TruB	Pseudouridine synthase	COG00130	pfam08068,pfam0150	cd02572	TIGR00425							1		1	1	1
arCOG04252	J	Translation, ribosomal structure and biogenesis	TruD	tRNA(Glu) U13 pseudouridine synthase TruD	COG00585	pfam01142	cd02577	TIGR00094	1	2	1	1		1	2		1	1	1
arCOG00761	J	Translation, ribosomal structure and biogenesis	TsaA	tRNA (Thr-GGU) A37 N-methylase	COG01720	pfam01980	cd09281	TIGR00104									1		
arCOG04733	J	Translation, ribosomal structure and biogenesis	Tsr3	Ribosome biogenesis protein Tsr3, contains Fer4-like metal-binding domain o	COG02042	pfam04034				1		1	1	1	1		1	1	2
arCOG01886	J	Translation, ribosomal structure and biogenesis	TyrS	Tyrosyl-tRNA synthetase	COG00162	pfam00579	cd00805	TIGR00234	1		1	1	1	1	1		1	1	2
arCOG04174	J	Translation, ribosomal structure and biogenesis	TYW1	Wybutosine (yW) biosynthesis enzyme, Fe-S oxidoreductase	COG00731	pfam04055,pfam0860	cd01335	TIGR03972	1	2	1				1	1	1	1	3
arCOG10124	J	Translation, ribosomal structure and biogenesis	TYW2	Wybutosine (yW) biosynthesis enzyme, TYW2 transferase	COG02520	pfam02475	cd02440	TIGR01444			1		1					1	2
arCOG04156	J	Translation, ribosomal structure and biogenesis	TYW3	Wybutosine (yW) biosynthesis enzyme	COG01590	pfam02676			1		1		1				1		
arCOG00808	J	Translation, ribosomal structure and biogenesis	ValS	Valyl-tRNA synthetase	COG00525	pfam00133,pfam0826	cd00817,cd079	TIGR00422	1	1	1					1	1	2	6
arCOG04225	J	Translation, ribosomal structure and biogenesis	YmdB	O-acetyl-ADP-ribose deacetylase (regulator of RNase III), contains Macro dom	COG02110	pfam01661	cd02907										1	1	1
arCOG00541	J	Translation, ribosomal structure and biogenesis	YSH1	Predicted exonuclease of the beta-lactamase fold involved in RNA processing	COG01236	pfam00753,pfam10996	pfam07521	TIGR03675	1			1			1	1	1	1	6
arCOG00545	J	Translation, ribosomal structure and biogenesis	YSH1	Predicted exonuclease of the beta-lactamase fold involved in RNA processing	COG01236			TIGR04122	1	1	1		1	2					
CELLULAR PROCESSES AND SIGNALING																			
arCOG02263	D	Cell cycle control, cell division, chromosome partit	-	Predicted cell division protein, SepF homolog	COG02450	pfam04472			1	1	1	1			1		1		
arCOG11012	D	Cell cycle control, cell division, chromosome partit	ATS1	Alpha-tubulin suppressor and related RCC1 domain-containing proteins	COG05184	pfam13540,pfam00415,pfam00415,pfam00415,pfam00415,pfa			1	2							7	3	6
arCOG04701	D	Cell cycle control, cell division, chromosome partit	CrcB/FEX	Integral membrane protein possibly involved in fluoride export and chromoso	COG00239	pfam02537		TIGR00494											3
arCOG02201	D	Cell cycle control, cell division, chromosome partit	FtsZ	Cell division GTPase	COG00206	pfam00091,pfam1232	cd02201	TIGR00065	2	3	1	2		1	1			2	1
arCOG05007	D	Cell cycle control, cell division, chromosome partit	Maf	Nucleotide-binding protein implicated in inhibition of septum formation	COG00424	pfam02545	cd00555	TIGR00172									2	1	3
arCOG03061	D	Cell cycle control, cell division, chromosome partit	MreB	Actin-like ATPase involved in cell morphogenesis	COG01077	pfam06723	cd10227	TIGR00904										1	1
arCOG00585	D	Cell cycle control, cell division, chromosome partit	Mrp	Mrp family protein, ATPase, contains iron-sulfur cluster	COG00489	pfam10609	cd02037	TIGR01969										1	3
arCOG00371	D	Cell cycle control, cell division, chromosome partit	Smc	Chromosome segregation ATPase	COG01196	pfam02463	cd03278,cd032	TIGR02169							1			3	
arCOG00586	D	Cell cycle control, cell division, chromosome partit	Soj	ATPase involved in chromosome partitioning, ParA family	COG01192	pfam01656	cd02042	TIGR03453											1
arCOG02416	N	Cell motility	-	Pilin/Flagellin, FlaG/FlaF family	COG03430	pfam07790				1					1			3	1
arCOG00434	N	Cell motility	-	AAA+ ATPase of MoxR-like family, a component of a putative secretion systen	COG00714	pfam07726	cd00009	TIGR02031	2	1	1	2			2	2	2	2	3
arCOG02079	N	Cell motility	-	S-layer protein, possibly associated with type IV pili like system	COG01361				1	1	1	1		1			1		
arCOG02423	N	Cell motility	-	Pilin/Flagellin, FlaG/FlaF family	COG03430	pfam07790			1	1	1		1	2			2		
arCOG02911	N	Cell motility	-	Predicted pilin/flagellin	SC.00343												1		
arCOG02382	N	Cell motility	CheB	Chemotaxis response regulator containing a CheY-like receiver domain and a	COG02201	pfam00072,pfam0133	cd00156	TIGR02875										1	
arCOG01819	N	Cell motility	CpaF	Pilus assembly protein, ATPase of CpaF family	COG04962	pfam00437	cd01130	TIGR03819	1	2	1	1	1	1		1		1	1
arCOG01829	N	Cell motility	FlaB	Archaeal flagellins	COG01681												3		1
arCOG02964	N	Cell motility	FlaD	Archaellum protein D/E	COG03351	pfam05377,pfam04659											2		
arCOG01824	N	Cell motility	FlaF	Archaellum protein F, flagellin of FlaG/FlaF family	COG03353												1		
arCOG01822	N	Cell motility	FlaG	Archaellum protein G, flagellin of FlaG/FlaF family	COG03354												1	1	
arCOG04148	N	Cell motility	FlaH	ATPase involved in biogenesis of archaellum	COG02874	pfam06745	cd01394	TIGR03881									1	1	
arCOG01809	N	Cell motility	FlaJ	Archaellum assembly protein J, TadC family	COG01955												1	1	
arCOG02298	N	Cell motility	FlaK/PulO	Peptidase A24A, prepilin type IV	COG01989	pfam01478,pfam06847			1	1	1	1		1	2		1	1	
arCOG01808	N	Cell motility	TadC	Pilus assembly protein TadC	COG02064	pfam00482,pfam00482			2	2	2	2	2	2	2		2	2	4
arCOG01811	N	Cell motility	TadC	Pilus assembly protein TadC	COG02064				1	1	1	1	1	1		1			
arCOG01812	N	Cell motility	TadC	Pilus assembly protein TadC	COG02064				1	3	1	1	1	1		1		1	1
arCOG01817	N	Cell motility	VirB11	ATPase involved in archaellum/pili biosynthesis	COG00630	pfam00437	cd01130	TIGR03819	1	1	1	1	1	1	1	1	3	2	2
arCOG02488	M	Cell wall/membrane/envelope biogenesis	-	Cell surface protein, a component of a putative secretion system	SC.00293				1	1	1	1			1	1	1		
arCOG01385	M	Cell wall/membrane/envelope biogenesis	-	Glycosyl transferase family 2	COG00463	pfam00535	cd02522	TIGR04283	1	1	1	1			1	1	2	2	3
arCOG05978	M	Cell wall/membrane/envelope biogenesis	-	Extracellular protein containing Kelch and FN3 domains		pfam07081				2					1	1	8		
arCOG00894	M	Cell wall/membrane/envelope biogenesis	-	Glycosyl transferase family 2	COG00463	pfam00535	cd04179	TIGR04182	2	1	2	2	2	1	2	1	1	2	3
arCOG02532	M	Cell wall/membrane/envelope biogenesis	-	Secreted Beta-propeller repeat protein fused to CARDB-like adhesion module	COG03291	pfam07705		TIGR04213	1			2	1	1	2	2	1	2	
arCOG07560	M	Cell wall/membrane/envelope biogenesis	-	Cell surface protein							1					1			1

arCOG03336	M	Cell wall/membrane/envelope biogenesis	-	Surface protein containing fasciclin-like repeats	COG02335	pfam02469			1		1		1						3	1
arCOG09173	M	Cell wall/membrane/envelope biogenesis	-	Cell surface protein		pfam07790,pfam0770	cd00146,cd001	TIGR00864,TIGR	1				1							
arCOG02086	M	Cell wall/membrane/envelope biogenesis	-	S-layer domain	COG01361				2											
arCOG03335	M	Cell wall/membrane/envelope biogenesis	-	Surface protein containing fasciclin-like repeats	COG02335	pfam02469				1										1
arCOG08795	M	Cell wall/membrane/envelope biogenesis	-	Predicted S-layer protein with Ig-like domain	SC.00192	pfam00127	cd13921	TIGR02657		1										
arCOG02482	M	Cell wall/membrane/envelope biogenesis	-	WD40/PQQ-like beta propeller repeat containing protein	COG01520	pfam13570,pfam1357	cd10276	TIGR03300			1									
arCOG01383	M	Cell wall/membrane/envelope biogenesis	-	Glycosyl transferase family 2	COG01216	pfam00535	cd04186	TIGR04017,TIGR01556						1				1	1	2
arCOG00568	M	Cell wall/membrane/envelope biogenesis	-	Oligosaccharyl transferase STT3 or related protein	COG01287													1		
arCOG01391	M	Cell wall/membrane/envelope biogenesis	-	Glycosyl transferase family 2	COG01215	pfam13641	cd06421	TIGR03030										1		
arCOG02080	M	Cell wall/membrane/envelope biogenesis	-	S-layer domain	COG01361													1		
arCOG02559	M	Cell wall/membrane/envelope biogenesis	-	Cell surface protein	SC.00325													1		
arCOG03553	M	Cell wall/membrane/envelope biogenesis	-	Cell surface protein	COG01572	pfam11824,pfam0770	5											1		
arCOG12808	M	Cell wall/membrane/envelope biogenesis	-	Glycosyltransferase, GT1 family	COG00438		cd03822											1		
arCOG01389	M	Cell wall/membrane/envelope biogenesis	-	Glycosyl transferase family 2	COG01215	pfam13641	cd06423	TIGR03937											1	1
arCOG00895	M	Cell wall/membrane/envelope biogenesis	-	Glycosyl transferase family 2	COG00463	pfam00535	cd06442	TIGR04182											1	
arCOG02487	M	Cell wall/membrane/envelope biogenesis	-	Cell surface protein, a component of a putative secretion system		pfam02369													1	
arCOG02538	M	Cell wall/membrane/envelope biogenesis	-	Secreted protein, with PKD repeat domain	COG03291			TIGR04213											1	
arCOG00896	M	Cell wall/membrane/envelope biogenesis	-	Glycosyl transferase family 2	COG00463	pfam00535	cd04179	TIGR04182											2	1
arCOG01410	M	Cell wall/membrane/envelope biogenesis	AglA	Glycosyltransferase	COG00438	pfam13579,pfam0053	cd03794	TIGR02149	5	5	3	3	1	1	2	4	1	1		
arCOG05365	M	Cell wall/membrane/envelope biogenesis	AglB	Oligosaccharyltransferase membrane subunit	COG01287	pfam02516,pfam1362	0,pfam13620,pf	TIGR04154	1	1	1	1	1	1	1	1	1	2	1	
arCOG00899	M	Cell wall/membrane/envelope biogenesis	AglD2	Predicted flippase	COG00392	pfam03706		TIGR00374	1		1						1			
arCOG00664	M	Cell wall/membrane/envelope biogenesis	AglF/RfbA	Glucose-1-phosphate uridyltransferase	COG01209	pfam00483	cd04181	TIGR03992									1	1		
arCOG03199	M	Cell wall/membrane/envelope biogenesis	AglH/Rfe	UDP-N-acetylmuramyl pentapeptide phosphotransferase/UDP-N-acetylglucos	COG00472	pfam00953	cd06856	TIGR00445	1								1			
arCOG01403	M	Cell wall/membrane/envelope biogenesis	AglL	Glycosyltransferase	COG00438	pfam13439,pfam0053	cd03804		3	2	1	2	1	1	1	3	6	5	9	
arCOG00253	M	Cell wall/membrane/envelope biogenesis	AglM/Ugd	UDP-glucose 6-dehydrogenase	COG01004	pfam03721,pfam0098	4,pfam03720	TIGR03026	1	1				1	1	2		1		
arCOG01381	M	Cell wall/membrane/envelope biogenesis	AglO	Glycosyl transferase family 2	COG00463	pfam00535	cd00761	TIGR04283										1	1	
arCOG02209	M	Cell wall/membrane/envelope biogenesis	AglR	MATE family membrane protein, Rfbx family	COG02244		cd13128		3	2	1	1	1	1	1	3	2			
arCOG10648	M	Cell wall/membrane/envelope biogenesis	BglC	Aryl-phospho-beta-D-glucosidase BglC, GH1 family	COG02730	pfam00150											1			
arCOG01375	M	Cell wall/membrane/envelope biogenesis	FlaA1	NDP-sugar epimerase, includes UDP-GlcNAc-inverting 4,6-dehydratase FlaA1	COG01086	pfam02719	cd05237	TIGR03589						1						
arCOG00665	M	Cell wall/membrane/envelope biogenesis	GalU	UDP-glucose pyrophosphorylase	COG01210	pfam00483	cd02541	TIGR01099	1					1	1			1	1	
arCOG00666	M	Cell wall/membrane/envelope biogenesis	GCD1	N-acetylglucosamine-1-phosphate uridyltransferase	COG01208	pfam00483,pfam0013	cd04181,cd056	TIGR03992	3	3	3	3	2		1	1	3	3	3	
arCOG00663	M	Cell wall/membrane/envelope biogenesis	GCD1	Nucleoside-diphosphate-sugar pyrophosphorylase involved in lipopolysaccha	COG01208	pfam00483	cd04181	TIGR03992							1					
arCOG00668	M	Cell wall/membrane/envelope biogenesis	GCD1	Nucleoside-diphosphate-sugar pyrophosphorylase involved in lipopolysaccha	COG01208	pfam00483,pfam0013	cd04181,cd056	TIGR03992	1	1										
arCOG00057	M	Cell wall/membrane/envelope biogenesis	GlmS	Glucosamine 6-phosphate synthetase	COG00449	pfam00310,pfam0138	cd00714,cd050	TIGR01135	1	1		1			1	1	1	1	1	
arCOG01373	M	Cell wall/membrane/envelope biogenesis	Gmd	GDP-D-mannose dehydratase	COG01089	pfam01370,pfam1395	cd05260	TIGR01472											1	
arCOG00068	M	Cell wall/membrane/envelope biogenesis	GutQ	Predicted sugar phosphate isomerase involved in capsule formation	COG00794	pfam01380	cd05005	TIGR03127									1			
arCOG02313	M	Cell wall/membrane/envelope biogenesis	LolE	ABC-type transport system, involved in lipoprotein release, permease compor	COG04591	pfam12704,pfam0268	7	TIGR02212											1	
arCOG06815	M	Cell wall/membrane/envelope biogenesis	MpaA	Murein tripeptide amidase MpaA	COG02866	pfam00246	cd06228												1	1
arCOG01568	M	Cell wall/membrane/envelope biogenesis	MscS	Small-conductance mechanosensitive channel	COG00668	pfam00924			2	2	1	1	1		1	1	1	2	5	
arCOG02821	M	Cell wall/membrane/envelope biogenesis	MurD	UDP-N-acetylmuramoylalanine-D-glutamate ligase	COG00771	pfam08245,pfam0287	5	TIGR01082					1	1	1					
arCOG02820	M	Cell wall/membrane/envelope biogenesis	MurE	UDP-N-acetylmuramyl tripeptide synthase	COG00769	pfam08245,pfam0287	5	TIGR01082	1	2								1		
arCOG07536	M	Cell wall/membrane/envelope biogenesis	NanM	N-acetylneuraminic acid mutarotase	COG03055													1		
arCOG02005	M	Cell wall/membrane/envelope biogenesis	RacX	Aspartate racemase	COG01794	pfam01177		TIGR00035									1			
arCOG01411	M	Cell wall/membrane/envelope biogenesis	RfaG	Glycosyltransferase	COG00438	pfam13439,pfam0053	cd03801	TIGR03999	2				1			1	1	1	3	
arCOG06759	M	Cell wall/membrane/envelope biogenesis	RfaG	Glycosyltransferase	COG00438					2										
arCOG01417	M	Cell wall/membrane/envelope biogenesis	RfaG	Glycosyltransferase	COG00438	pfam00534	cd01635										1			
arCOG01405	M	Cell wall/membrane/envelope biogenesis	RfaG	Glycosyltransferase	COG00438	pfam13439,pfam1369	cd03794	TIGR04063											1	
arCOG01409	M	Cell wall/membrane/envelope biogenesis	RfaG	Glycosyltransferase	COG00438	pfam13692	cd03794	TIGR03087											1	
arCOG01371	M	Cell wall/membrane/envelope biogenesis	RfbB	dTDP-D-glucose 4,6-dehydratase	COG01088	pfam01370	cd05246	TIGR01181									1			
arCOG04188	M	Cell wall/membrane/envelope biogenesis	RfbC	dTDP-4-dehydrorhamnose 3,5-epimerase or related enzyme	COG01898	pfam00908		TIGR01221									1			
arCOG01367	M	Cell wall/membrane/envelope biogenesis	RfbD	dTDP-4-dehydrorhamnose reductase	COG01091	pfam04321	cd05254	TIGR01214									1			
arCOG01168	M	Cell wall/membrane/envelope biogenesis	RspA	L-alanine-DL-glutamate epimerase or related enzyme of enolase superfamily	COG04948	pfam02746,pfam0118	cd03316	TIGR02534									1			
arCOG04827	M	Cell wall/membrane/envelope biogenesis	TagB	Glycosyl/glycerophosphate transferase involved in teichoic acid biosynthesis	COG01887	pfam04464						1					1			
arCOG01222	M	Cell wall/membrane/envelope biogenesis	TagD	Cytidylyltransferase fused to conserved domain of DUF357 family	COG00615	pfam01467	cd02170	TIGR02199	2	1	1	2		1	1		1	1	3	
arCOG04468	M	Cell wall/membrane/envelope biogenesis	WcaG	Nucleoside-diphosphate-sugar epimerase	COG00451	pfam01370,pfam1395	cd05272	TIGR01179	1	2					1		1		3	
arCOG01369	M	Cell wall/membrane/envelope biogenesis	WcaG	Nucleoside-diphosphate-sugar epimerase	COG00451	pfam01370	cd05234	TIGR01179	4	4	1			2	3	1	2	6	4	
arCOG04826	M	Cell wall/membrane/envelope biogenesis	WcaK	Polysaccharide pyruvyl transferase family protein	COG02327	pfam04230		TIGR03609									1			
arCOG01392	M	Cell wall/membrane/envelope biogenesis	WecB	UDP-N-acetylglucosamine 2-epimerase	COG00381	pfam02350	cd03786	TIGR00236				1			1				1	
arCOG00252	M	Cell wall/membrane/envelope biogenesis	WecC	UDP-N-acetyl-D-mannosaminuronate dehydrogenase	COG00677	pfam03721,pfam0098	cd05266	TIGR03026	1	1	1	1					1	1	1	
arCOG00118	M	Cell wall/membrane/envelope biogenesis	WecE	Predicted pyridoxal phosphate-dependent enzyme apparently involved in reg	COG00399	pfam01041	cd00616	TIGR03588	3	1	1	1		2				2	1	
arCOG03634	M	Cell wall/membrane/envelope biogenesis	WecH	Surface polysaccharide O-acyltransferase, integral membrane enzyme	COG03274	pfam01757			1	1	1									
arCOG03787	V	Defense mechanisms	-	LEA14-like dessication related protein	COG05608	pfam03168,pfam0316	8		1	1		2			1					
arCOG08906	V	Defense mechanisms	-	Predicted antitoxins containing the HTH domain	COG02886	pfam03683									1					
arCOG01467	V	Defense mechanisms	-	ABC-type multidrug transport system, permease component	COG00842	pfam01061		TIGR01247	2	1	1				2	1				
arCOG03966	V	Defense mechanisms	-	CopG family DNA-binding protein												1				
arCOG01463	V	Defense mechanisms	-	ABC-type multidrug transport system, permease component	COG00842	pfam01061		TIGR01247	1	1	1	1		1						
arCOG00722	V	Defense mechanisms	-	Predicted antitoxins containing the HTH domain	COG02886	pfam03683			1	2	1	1	1	1						
arCOG04735	V	Defense mechanisms	-	CopG/MetJ, RHH domain containing DNA-binding protein, often an antitoxin i	COG03609					1									1	1
arCOG00713	V	Defense mechanisms	-	PIN domain containing protein	COG01848	pfam01850	cd09861										1			
arCOG01009	V	Defense mechanisms	-	CopG/MetJ, RHH domain containing DNA-binding protein, often an antitoxin i	COG03609												1			
arCOG01469	V	Defense mechanisms	-	ABC-type multidrug transport system, permease component	COG00842												1			
arCOG02123	V	Defense mechanisms	-	HEPN domain containing protein	COG01895	pfam05168											1			
arCOG02780	V	Defense mechanisms	-	Endonuclease, HJR/Mrr/RecB family													1			
arCOG05724	V	Defense mechanisms	-	RecB family restriction endonuclease, SeqA-like protein	COG02810	pfam04313,pfam0392	5										1			
arCOG06522	V	Defense mechanisms	-	HNH nuclease domain containing protein	COG03183												1			
arCOG06948	V	Defense mechanisms	-	PIN domain	SC.00610												1			
arCOG06971	V	Defense mechanisms	-	CopG/RHH family DNA binding protein	COG03609												1			
arCOG09828	V	Defense mechanisms	-	RelB family antitoxin	SC.00879												1			
arCOG06966	V	Defense mechanisms	-	CopG/RHH family DNA binding protein	COG03609												2			
arCOG03166	V	Defense mechanisms	-	AAA+ superfamily ATPase fused to HTH and RecB nuclease domains	COG01672												4	1	5	
arCOG03521	V	Defense mechanisms	-	Type II restriction enzyme, methylase subunit	COG01002	pfam12950												1		

arCOG00815	V	Defense mechanisms	AbrB	AbrB family transcriptional regulator	COG02002	pfam04014		TIGR01439									1			
arCOG01452	V	Defense mechanisms	Cas1	CRISPR-associated protein Cas1	COG01518	pfam01867	cd09636	TIGR00287									2			
arCOG02666	V	Defense mechanisms	Cas10	CRISPR associated protein Cas10, large subunit of Type III system effector com	COG01353	pfam01966											1			
arCOG04194	V	Defense mechanisms	Cas2	CRISPR-associated protein Cas2	COG01343	pfam09827	cd09725	TIGR01573									1			
arCOG00786	V	Defense mechanisms	Cas4	CRISPR-associated protein Cas4, RecB family exonuclease	COG01468	pfam01930	cd09637	TIGR00372	1	2	1	1		1	1		1	1		
arCOG00790	V	Defense mechanisms	Cas4	CRISPR-associated protein Cas4, RecB family exonuclease	COG01468		cd09637	TIGR00372									1			
arCOG00794	V	Defense mechanisms	Cas4	CRISPR-associated protein Cas4, RecB family nuclease	COG01468	pfam01930	cd09637	TIGR00372									1			
arCOG04342	V	Defense mechanisms	Cas6	CRISPR-Cas system related protein Cas6, RAMP superfamily	COG01583	pfam01881	cd09652	TIGR01877									1			
arCOG00196	V	Defense mechanisms	CcmA	ABC-type multidrug transport system, ATPase component	COG01131	pfam00005	cd03230	TIGR03740	2	1	1			1	1	1		1	2	
arCOG00194	V	Defense mechanisms	CcmA	ABC-type multidrug transport system, ATPase component	COG01131	pfam00005	cd03230	TIGR01188	7	7	6	2	1	1	6	3	5	4	6	
arCOG02665	V	Defense mechanisms	Cmr3g5	CRISPR-Cas system related protein, RAMP superfamily Cas5 group	COG01769	pfam09700											1			
arCOG06487	V	Defense mechanisms	Csm2SS	CRISPR/Cas system CSM-associated protein Csm2, small subunit	COG01421	pfam03750	cd09647	TIGR01870									1			
arCOG02658	V	Defense mechanisms	Csm3g7	CRISPR-Cas system related protein, RAMP superfamily Cas7 group	COG01337	pfam03787	cd09683	TIGR02581									1			
arCOG03222	V	Defense mechanisms	Csm4g5	CRISPR-Cas system related protein, RAMP superfamily Cas5 group	COG01567		cd09663	TIGR01903									1			
arCOG03718	V	Defense mechanisms	Csm5g7	CRISPR/Cas system CSM-associated protein Csm5, group 7 of RAMP superfam	COG01332		cd09662	TIGR01899									1			
arCOG01449	V	Defense mechanisms	Csx1	CARF domain containing protein	COG01517			TIGR01884									1			
arCOG07641	V	Defense mechanisms	Csx1	CARF domain containing protein, contains HTH and HEPN domains	COG01517	pfam09455	cd09732	TIGR02221									1			
arCOG03847	V	Defense mechanisms	Csx1	CARF domain containing protein	COG01517												2			
arCOG00373	V	Defense mechanisms	DndD	DNA sulfur modification protein DndD, ATPase	COG01196	pfam13514		TIGR02169								1	1			
arCOG02632	V	Defense mechanisms	HsdM	Type I restriction-modification system methyltransferase subunit	COG00286	pfam12161,pfam02384		TIGR00497									1		2	
arCOG03779	V	Defense mechanisms	McrB	GTPase subunit of restriction endonuclease	COG01401	pfam07728	cd00009		2	1	1					2	1			
arCOG05102	V	Defense mechanisms	McrC	McrBC 5-methylcytosine restriction system component	COG04268	pfam10117											1			
arCOG02841	V	Defense mechanisms	MdlB	ABC-type multidrug transport system, ATPase and permease component	COG01132	pfam00664,pfam0000	cd03253	TIGR02203	1	1		1	1				2	2	3	
arCOG02777	V	Defense mechanisms	Mrr	Restriction endonuclease	COG01715	pfam04471												1		
arCOG01008	V	Defense mechanisms	NikR	Transcriptional regulator, CopG/Arc/MetJ family (DNA-binding and a metal-bi	COG00864	pfam08753		TIGR02793									1	1	1	
arCOG02836	V	Defense mechanisms	NikR	Transcriptional regulator, CopG/Arc/MetJ family (DNA-binding and a metal-bi	COG00864												1			
arCOG00516	V	Defense mechanisms	NimA	Nitroimidazol reductase NimA or a related FMN-containing flavoprotein, pyrid	COG03467	pfam01243		TIGR04023	1			1								
arCOG00525	V	Defense mechanisms	NimA	Nitroimidazol reductase NimA or a related FMN-containing flavoprotein, pyrid	COG03467	pfam12724,pfam01243				2	1			1						
arCOG01731	V	Defense mechanisms	NorM	Na+-driven multidrug efflux pump	COG00534	pfam01554,pfam0155	cd13137	TIGR00797	1	1	1	1	1				3			
arCOG01663	V	Defense mechanisms	RelE	Cytotoxic translational repressor of toxin-antitoxin stability system	COG02026	pfam05016										1				
arCOG01665	V	Defense mechanisms	RelE	Cytotoxic translational repressor of toxin-antitoxin stability system	COG02026	pfam05016											3			
arCOG00922	V	Defense mechanisms	SalX	ABC-type antimicrobial peptide transport system, ATPase component	COG01136	pfam00005	cd03255	TIGR02673	3	2	3	1			2	2	2	4	7	
arCOG02312	V	Defense mechanisms	SalY	ABC-type antimicrobial peptide transport system, permease component	COG00577	pfam02687		TIGR02212	1	1	1	1					2	2	3	
arCOG02957	U	Intracellular trafficking, secretion, and vesicular tr	SBH1	Preprotein translocase subunit Sec61beta	COG04023	pfam03911			1	1					1	1	1		2	
arCOG02673	U	Intracellular trafficking, secretion, and vesicular tr	-	OxaA/SpoJ/YidC translocase/secretase, sec-independent itegration of nascent	COG01422	pfam01956			1	2				1	1	1	1	1	1	
arCOG04169	U	Intracellular trafficking, secretion, and vesicular tr	SecY	Preprotein translocase subunit SecY	COG00201	pfam10559,pfam00344		TIGR00967	1	2				1	1	1	1	1	2	
arCOG01217	U	Intracellular trafficking, secretion, and vesicular tr	SEC65	Signal recognition particle 19 kDa protein	COG01400				1		1				1	1	1	1	2	
arCOG04736	U	Intracellular trafficking, secretion, and vesicular tr	TatC	Sec-independent protein secretion pathway component TatC	COG00805	pfam00902,pfam00902		TIGR01912,TIGR01912						1	1	1			1	
arCOG01739	U	Intracellular trafficking, secretion, and vesicular tr	LepB	Signal peptidase I	COG00681	pfam00717	cd06530	TIGR02228	1	1	1	1	1	1	1		1	1	2	
arCOG02204	U	Intracellular trafficking, secretion, and vesicular tr	Sss1	Preprotein translocase subunit Sss1	COG02443				1	2	1	1		1	1		1	1	2	
arCOG01919	U	Intracellular trafficking, secretion, and vesicular tr	TatC	Sec-independent protein secretion pathway component TatC	COG00805	pfam00902		TIGR00945	1	2		1			1			2	3	
arCOG01228	U	Intracellular trafficking, secretion, and vesicular tr	Ffh	Signal recognition particle GTPase	COG00541	pfam02881,pfam0044	cd03115	TIGR00959		1	1	1			1		1	1	1	
arCOG02694	U	Intracellular trafficking, secretion, and vesicular tr	TatA	Sec-independent protein secretion pathway component	COG01826			TIGR01411		2	1			1	1				1	
arCOG04471	U	Intracellular trafficking, secretion, and vesicular tr	-	Exosortase	COG04083	pfam09721		TIGR04125							1		1	1	1	
arCOG01227	U	Intracellular trafficking, secretion, and vesicular tr	FtsY	Signal recognition particle GTPase	COG00552	pfam02881,pfam0044	cd03115	TIGR00064	1								1	1	5	
arCOG01997	U	Intracellular trafficking, secretion, and vesicular tr	MarC	Multiple antibiotic transporter	COG02095	pfam01914		TIGR00427									1			
arCOG03382	U	Intracellular trafficking, secretion, and vesicular tr	TolB	Periplasmic component of the Tol biopolymer transport system	COG00823	pfam00930		TIGR02800									1			
arCOG04816	U	Intracellular trafficking, secretion, and vesicular tr	-	TraG/TraD/VirD4 family enzyme, ATPase	COG03505	pfam12846											1			
arCOG07496	U	Intracellular trafficking, secretion, and vesicular tr	VirB4	Type IV secretory pathway, VirB4 component	COG03451												1			
arCOG03383	U	Intracellular trafficking, secretion, and vesicular tr	TolB	Periplasmic component of the Tol biopolymer transport system	COG00823	pfam00930,pfam08308		TIGR02800										1		
arCOG01241	X	Mobilome: prophages, transposons	XerC	XerD/XerC family integrase	COG00582	pfam13495,pfam0058	cd00796	TIGR02224								1	4	1	2	
arCOG03473	X	Mobilome: prophages, transposons	-	Transposase	COG05421					2										
arCOG03989	X	Mobilome: prophages, transposons	XerC	XerD/XerC family integrase	COG00582	pfam00589	cd00397										1			
arCOG10214	X	Mobilome: prophages, transposons	-	Relaxase/mobilization nuclease domain-containing protein													1			
arCOG10424	X	Mobilome: prophages, transposons	GepA	Uncharacterized phage-associated protein	COG03600												1			
arCOG02751	X	Mobilome: prophages, transposons	-	Transposase, IS5 family	COG03039	pfam01609											10			
arCOG01915	O	Posttranslational modification, protein turnover, c	HflC	Membrane protease subunit, stomatin/prohibitin homolog	COG00330	pfam01145	cd08826	TIGR01933	1	1	1				1	1	1		1	
arCOG01912	O	Posttranslational modification, protein turnover, c	-	Membrane protein implicated in regulation of membrane protease activity	COG01585	pfam01957			1	1	1				1	1	1			
arCOG01306	O	Posttranslational modification, protein turnover, c	RPT1	ATP-dependent 26S proteasome regulatory subunit	COG01222	pfam00004	cd00009	TIGR01242	1	1		1			1	1	1	2	3	
arCOG06181	O	Posttranslational modification, protein turnover, c	TrxA	Thiol-disulfide isomerase or thioredoxin	COG00526	pfam00578	cd02966		1	1			1		1	1				
arCOG01341	O	Posttranslational modification, protein turnover, c	GIM5	Predicted prefoldin, molecular chaperone implicated in de novo protein foldir	COG01730	pfam02996	cd00584	TIGR00293	1	1					1	1	1			
arCOG00609	O	Posttranslational modification, protein turnover, c	RseP	Membrane-associated protease RseP, regulator of RpoE activity in bacteria	COG00750	pfam02163	cd06160		1	2	1	1	1	1	1	1	1			
arCOG02959	O	Posttranslational modification, protein turnover, c	Iap	Zn-dependent amino- or carboxypeptidase, M28 family	COG02234	pfam04389	cd05643		1	2					1	1		3		
arCOG00314	O	Posttranslational modification, protein turnover, c	TrxA	Thiol-disulfide isomerase or thioredoxin	COG00526	pfam00578	cd02969		1		1	1	1	1	1	1		1	2	
arCOG04463	O	Posttranslational modification, protein turnover, c	-	Presenilin-like membrane protease, A22 family	COG03389	pfam06550			1		1	1	1		1	1	1	1	1	
arCOG01257	O	Posttranslational modification, protein turnover, c	GroEL	Chaperonin GroEL, HSP60 family	COG00459	pfam00118	cd03343	TIGR02339	1		1	1		1	1	1	2	3	9	
arCOG01833	O	Posttranslational modification, protein turnover, c	IbpA	Molecular chaperone (HSP20 family)	COG00071	pfam00011	cd06464		1			1		1	1	1	1			
arCOG01296	O	Posttranslational modification, protein turnover, c	TrxB	Thioredoxin reductase	COG00492	pfam07992		TIGR01292	2	1	1		1	2	1	1	2	1	4	
arCOG04767	O	Posttranslational modification, protein turnover, c	PpiB	Peptidyl-prolyl cis-trans isomerase (rotamase) - cyclophilin family	COG00652				2	2	1	2	1	1	1	1		3	3	
arCOG06880	O	Posttranslational modification, protein turnover, c	DnaJ	DnaJ type Zn finger domain	COG00484			TIGR02349	2	2	2		1		1	1	1	1	4	
arCOG02553	O	Posttranslational modification, protein turnover, c	AprE	Subtilisin-like serine protease	COG01404	pfam05048				2	1	1	1		1	1				
arCOG00980	O	Posttranslational modification, protein turnover, c	SlpA	FKBP-type peptidyl-prolyl cis-trans isomerase 2	COG01047	pfam00254									1	1	1	1	2	
arCOG01139	O	Posttranslational modification, protein turnover, c	Rri1	Proteasome lid subunit RPN8/RPN11, contains Jab1/MPN domain metalloenz	COG01310	pfam14464	cd08072								1	1	1			
arCOG03610	O	Posttranslational modification, protein turnover, c	AprE	Subtilisin-like serine protease	COG01404		cd02619			1			1		1	2			3	
arCOG01331	O	Posttranslational modification, protein turnover, c	HtpX	Zn-dependent protease with chaperone function	COG00501	pfam01435			1	1	1	1		1	1					
arCOG00971	O	Posttranslational modification, protein turnover, c	PRE1	20S proteasome, alpha subunit	COG00638	pfam10584,pfam0022	cd03756	TIGR03633	1	1	1		1		1		1	1	1	
arCOG02436	O	Posttranslational modification, protein turnover, c	NosY	ABC-type transport system involved in multi-copper enzyme maturation, perr	COG01277	pfam12679			1	1	1				1					
arCOG05153	O	Posttranslational modification, protein turnover, c	GroS	Co-chaperonin GroES (HSP10)	COG00234	pfam00166	cd00320		1	1					1					
arCOG05154	O	Posttranslational modification, protein turnover, c	GroEL	Chaperonin GroEL (HSP60 family)	COG00459	pfam00118	cd03344	TIGR02348	1	1					1					
arCOG02816	O	Posttranslational modification, protein turnover, c	MsrA	Peptide methionine sulfoxide reductase	COG00225	pfam01625		TIGR00401	1	2	1				1			1	3	
arCOG00970	O	Posttranslational modification, protein turnover, c	PRE1	20S proteasome, beta subunit	COG00638	pfam00227	cd03764	TIGR03634	1		1	1	1	1	1		1	1	5	

arCOG00270	O	Posttranslational modification, protein turnover, c	CcmF	Cytochrome c biogenesis factor	COG01138	pfam01578		TIGR00353	1			1			1			1	2
arCOG00310	O	Posttranslational modification, protein turnover, c	Bcp	Peroxiredoxin	COG01225	pfam00578	cd03018	TIGR03137	2	2					1		1	2	11
arCOG04560	O	Posttranslational modification, protein turnover, c	IscA	Fe-S cluster assembly iron-binding protein IscA	COG00316	pfam01521		TIGR00049	2	3		1	1		1			1	4
arCOG03607	O	Posttranslational modification, protein turnover, c	-	Cysteine protease, C1A family	COG04870	pfam00112	cd02619			1	1				1		1		
arCOG01845	O	Posttranslational modification, protein turnover, c	PaaD	Metal-sulfur cluster biosynthetic enzyme	COG02151	pfam01883		TIGR02945		1	1				1		2	1	1
arCOG04236	O	Posttranslational modification, protein turnover, c	SufC	Cysteine desulfurase activator ATPase	COG00396	pfam00005	cd03217	TIGR01978		1	1				1		2	1	1
arCOG04064	O	Posttranslational modification, protein turnover, c	RseP	Membrane-associated protease RseP, regulator of RpoE activity in bacteria	COG00750	pfam02163	cd06159	TIGR00054		1					1		1		2
arCOG01226	O	Posttranslational modification, protein turnover, c	ArgK	Putative periplasmic protein kinase ArgK or related GTPase of G3E family	COG01703	pfam03308	cd03114	TIGR00750							1		1	1	2
arCOG01308	O	Posttranslational modification, protein turnover, c	Cdc48	ATPase of the AAA+ class , CDC48 family	COG00464	pfam00004,pfam0000	cd00009,cd000	TIGR01243							1		2	2	4
arCOG00267	O	Posttranslational modification, protein turnover, c	CcmC	ABC-type transport system involved in cytochrome c biogenesis, permease co	COG00755	pfam01578		TIGR01191							1			1	1
arCOG02607	O	Posttranslational modification, protein turnover, c	GrxC	Glutaredoxin	COG00695	pfam00462	cd02976	TIGR02196							1			1	1
arCOG04957	O	Posttranslational modification, protein turnover, c	CcmE	Cytochrome c-type biogenesis protein CcmE	COG02332										1			1	3
arCOG02438	O	Posttranslational modification, protein turnover, c	NosY	ABC-type transport system involved in multi-copper enzyme maturation, perr	COG01277	pfam12679									1				
arCOG03614	O	Posttranslational modification, protein turnover, c	-	Cysteine protease, C1A family	COG04870	pfam00112	cd02619								1				
arCOG00065	O	Posttranslational modification, protein turnover, c	csdA	Selenocysteine lyase/Cysteine desulfurase	COG00520	pfam00266	cd06453	TIGR01979	3	5	4	1		1	2	1		4	2
arCOG00479	O	Posttranslational modification, protein turnover, c	CyoE	Polyprenyltransferase (cytochrome oxidase assembly factor)	COG00109	pfam01040	cd13957	TIGR01473		1					2	1		1	1
arCOG04142	O	Posttranslational modification, protein turnover, c	DYS1	Deoxyhypusine synthase	COG01899	pfam01916		TIGR00321	2	2	2	2	1	2	2	2	1	2	4
arCOG03669	O	Posttranslational modification, protein turnover, c	-	Subtilase family protease	COG04934	pfam09286	cd11377,cd04056		4	1	1	3	1	2	2	3	1		3
arCOG04772	O	Posttranslational modification, protein turnover, c	GrpE	Molecular chaperone GrpE (heat shock protein)	COG00576	pfam01025	cd00446		1	2	1			1	2		1	1	2
arCOG03060	O	Posttranslational modification, protein turnover, c	DnaK	Chaperone DnaK (HSP70)	COG00443	pfam00012	cd10234	TIGR02350	1	2	1			1	2		1	1	3
arCOG03103	O	Posttranslational modification, protein turnover, c	CtaA	Uncharacterized protein required for cytochrome oxidase assembly	COG01612	pfam02628			2	1	2		2		2			1	1
arCOG01972	O	Posttranslational modification, protein turnover, c	TrxA	Thiol-disulfide isomerase or thioredoxin	COG00526	pfam00085	cd02947	TIGR01068	2	2	1	1		1	2			2	3
arCOG01715	O	Posttranslational modification, protein turnover, c	SufB	Cysteine desulfurase activator SufB	COG00719	pfam01458		TIGR01981		2	2				2		1	2	2
arCOG00702	O	Posttranslational modification, protein turnover, c	AprE	Subtilisin-like serine protease	COG01404	pfam00082	cd07477	TIGR03921	5	6	4	4	4	1	5	2		3	7
arCOG02007	O	Posttranslational modification, protein turnover, c	-	Prenyltransferase family protein containing thioredoxin domain	COG01331	pfam03190	cd02955			1				1		1			
arCOG02834	O	Posttranslational modification, protein turnover, c	RseP	Membrane-associated protease RseP, regulator of RpoE activity in bacteria	COG00750	pfam02163,pfam1318	cd06159,cd009	TIGR00054						1		1		1	
arCOG00976	O	Posttranslational modification, protein turnover, c	Pcm	Protein-L-isoaspartate carboxylmethyltransferase	COG02518	pfam01135	cd02440	TIGR00080								1	1	1	2
arCOG03202	O	Posttranslational modification, protein turnover, c	-	Collagenase family protease	COG00826	pfam01297,pfam01136										1			
arCOG06807	O	Posttranslational modification, protein turnover, c	-	Cytochrome c biogenesis factor	COG01138											1			
arCOG02160	O	Posttranslational modification, protein turnover, c	LonB	Predicted ATP-dependent protease	COG01067	pfam01078,pfam13654	pfam05362	TIGR00764	1	1	1	1					1	1	2
arCOG00312	O	Posttranslational modification, protein turnover, c	AhpC	AhpC/TSA family peroxiredoxin	COG00450	pfam00578,pfam1041	cd03016	TIGR03137	1	1	1	1					1	1	
arCOG01342	O	Posttranslational modification, protein turnover, c	GimC	Prefoldin, chaperonin cofactor	COG01382	pfam01920	cd00632	TIGR02338	1	1							1		1
arCOG01328	O	Posttranslational modification, protein turnover, c	CcmB	ABC-type transport system involved in cytochrome c biogenesis, permease co	COG02386				1	1									
arCOG01846	O	Posttranslational modification, protein turnover, c	PaaD	Metal-sulfur cluster biosynthetic enzyme	COG02151	pfam01883,pfam0914	cd02037	TIGR02945,TIGR	1	1									
arCOG02815	O	Posttranslational modification, protein turnover, c	-	Conserved domain frequently associated with peptide methionine sulfoxide r	COG00229	pfam01641		TIGR00357	1	2									
arCOG00981	O	Posttranslational modification, protein turnover, c	SlpA	FKBP-type peptidyl-prolyl cis-trans isomerase 2	COG01047	pfam00254		TIGR00115	1			1					1	1	1
arCOG03580	O	Posttranslational modification, protein turnover, c	STE14	Putative protein-S-isoprenylcysteine methyltransferase	COG02020	pfam04191			1			1					1		
arCOG06823	O	Posttranslational modification, protein turnover, c	AprE	Subtilisin-like serine protease	COG01404	pfam00082	cd05562	TIGR03921	1									1	1
arCOG02173	O	Posttranslational modification, protein turnover, c	NrdG	Organic radical activating enzyme	COG00602	pfam13394	cd01335	TIGR03963		1							1	1	1
arCOG02962	O	Posttranslational modification, protein turnover, c	Iap	Zn-dependent amino- or carboxypeptidase, M28 family	COG02234	pfam04389	cd02690			1								1	
arCOG02443	O	Posttranslational modification, protein turnover, c	NosY	ABC-type transport system involved in multi-copper enzyme maturation, perr	COG01277	pfam12679				1									
arCOG03028	O	Posttranslational modification, protein turnover, c	NifU	Fe-S cluster biogenesis protein NfuA, 4Fe-4S-binding domain	COG00694	pfam01106		TIGR02000		1									
arCOG01297	O	Posttranslational modification, protein turnover, c	AhpF	Alkyl hydroperoxide reductase, large subunit	COG03634	pfam07992		TIGR03140			1								
arCOG02606	O	Posttranslational modification, protein turnover, c	GrxC	Glutaredoxin	COG00695	pfam00462	cd02976	TIGR02196				1							
arCOG10598	O	Posttranslational modification, protein turnover, c	-	ATP-dependent protease La PUA-like domain		pfam02190								1				1	1
arCOG00185	O	Posttranslational modification, protein turnover, c	GsiA	ABC-type glutathione transport system ATPase component, contains duplicate	COG01123	pfam00005,pfam0000	cd03225,cd032	TIGR03269									1	1	1
arCOG05850	O	Posttranslational modification, protein turnover, c	Pcp	Pyrrolidone-carboxylate peptidase (N-terminal pyroglutamyl peptidase)	COG02039	pfam01470	cd00501	TIGR00504									1	1	
arCOG00164	O	Posttranslational modification, protein turnover, c	CysU	ABC-type sulfate transport system, permease component	COG00555		cd06261	TIGR02141									1		
arCOG00299	O	Posttranslational modification, protein turnover, c	-	DsbA family protein													1		
arCOG00636	O	Posttranslational modification, protein turnover, c	HypE	Hydrogenase maturation factor	COG00309	pfam00586,pfam0276	cd06061	TIGR02124									1		
arCOG00946	O	Posttranslational modification, protein turnover, c	PflA	Pyruvate-formate lyase-activating enzyme	COG01180	pfam04055	cd01335	TIGR04337									1		
arCOG01187	O	Posttranslational modification, protein turnover, c	HypF	Hydrogenase maturation factor	COG00068	pfam00708,pfam07503	pfam07503,pf	TIGR00143									1		
arCOG01218	O	Posttranslational modification, protein turnover, c	TrxA	Thiol-disulfide isomerase or thioredoxin	COG00526	pfam08484,pfam1319	cd02975,cd029	TIGR02187									1		
arCOG01716	O	Posttranslational modification, protein turnover, c	SufB	Cysteine desulfurase activator SufB	COG00719	pfam01458		TIGR01980									1		
arCOG01832	O	Posttranslational modification, protein turnover, c	IbpA	Molecular chaperone (HSP20 family)	COG00071	pfam00011	cd06464										1		
arCOG02162	O	Posttranslational modification, protein turnover, c	LonB	Predicted ATP-dependent protease	COG01067	pfam01078,pfam1324	cd00009,cd000	TIGR00764									1		
arCOG02737	O	Posttranslational modification, protein turnover, c	-	Predicted Fe-Mo cluster-binding protein, NifX family	COG01433	pfam02579	cd00562										1		
arCOG03689	O	Posttranslational modification, protein turnover, c	NosY	ABC-type transport system involved in multi-copper enzyme maturation, perr	COG01277												1		
arCOG04427	O	Posttranslational modification, protein turnover, c	HypC	Hydrogenase maturation factor	COG00298	pfam01455		TIGR00074									1		
arCOG04428	O	Posttranslational modification, protein turnover, c	HypD	Hydrogenase maturation factor	COG00409	pfam01924		TIGR00075									1		
arCOG00322	O	Posttranslational modification, protein turnover, c	TldE	Inactivated Zn-dependent protease, component of TldD/TldE system	COG00312	pfam01523											2	1	1
arCOG00321	O	Posttranslational modification, protein turnover, c	TldD	Zn-dependent protease, component of TldD/TldE system	COG00312	pfam01523											2	1	
arCOG03685	O	Posttranslational modification, protein turnover, c	-	Predicted redox protein, regulator of disulfide bond formation	COG01765	pfam02566											2	1	
arCOG00952	O	Posttranslational modification, protein turnover, c	PflA	Pyruvate-formate lyase-activating enzyme	COG01180	pfam04055		TIGR02494									2		
arCOG04933	O	Posttranslational modification, protein turnover, c	-	Serine protease inhibitor	COG04826	pfam00079	cd00172										2		
arCOG02734	O	Posttranslational modification, protein turnover, c	-	Predicted Fe-Mo cluster-binding protein, NifX family	COG01433	pfam02579	cd00851										4		
arCOG07441	O	Posttranslational modification, protein turnover, c	SurA	Parvulin-like peptidyl-prolyl isomerase	COG00760	pfam00639		TIGR02933										1	1
arCOG01929	O	Posttranslational modification, protein turnover, c	XdhC	Xanthine and CO dehydrogenase maturation factor, XdhC/CoxF family	COG01975	pfam13478												1	2
arCOG04648	O	Posttranslational modification, protein turnover, c	-	Glutaredoxin-related protein	COG00278	pfam00462	cd03028	TIGR00365										1	2
arCOG02784	O	Posttranslational modification, protein turnover, c	SRAP	Putative SOS response-associated peptidase YedK	COG02135	pfam02586												1	
arCOG02846	O	Posttranslational modification, protein turnover, c	DnaJ	DnaJ-class molecular chaperone	COG00484	pfam00226	cd06257	TIGR02349										1	
arCOG03947	O	Posttranslational modification, protein turnover, c	clpA	ATP-binding subunits of Clp protease and DnaK/DnaJ chaperones	COG00542	pfam02861,pfam0286	cd00009,cd000	TIGR03346										1	
arCOG04712	O	Posttranslational modification, protein turnover, c	ECM4	Predicted glutathione S-transferase	COG00435	pfam13409,pfam1341	cd03190											1	
arCOG03026	O	Posttranslational modification, protein turnover, c	NifU	Fe-S cluster biogenesis protein NfuA, 4Fe-4S-binding domain	COG00694	pfam01106		TIGR02000											1
arCOG01153	T	Signal transduction mechanisms	CpdA	3',5'-cyclic AMP phosphodiesterase CpdA	COG01409	pfam00149,pfam0145	cd07400,cd001	TIGR00040,TIGR02689		1					1	1			
arCOG02985	T	Signal transduction mechanisms	-	Membrane proteinase of CAAX superfamily, regulator of anti-sigma factor	COG02339	pfam13367			2	2	1	1	1		2	1	1	1	2
arCOG04425	T	Signal transduction mechanisms	Wzb	Protein-tyrosine-phosphatase	COG00394	pfam01451	cd00115	TIGR02689		2				1	2	1		1	4
arCOG01171	T	Signal transduction mechanisms	RAD55	RecA-superfamily ATPase implicated in signal transduction	COG00467	pfam06745	cd01124	TIGR03877	2	2					2	2	10	3	
arCOG02967	T	Signal transduction mechanisms	-	NACHT family NTPase fused to HEAT repeats domain	COG05635	pfam13646,pfam13646			1	1	1	2				1		1	
arCOG02053	T	Signal transduction mechanisms	UspA	Nucleotide-binding protein, UspA family	COG00589	pfam00582	cd00293		1							1	1		
arCOG05374	T	Signal transduction mechanisms	-	GAF domain-containing protein	COG01956	pfam13185			1							1			

arCOG06712	T	Signal transduction mechanisms	AtoS	Sensory protein, contains PAS domain	COG02202	pfam13426	cd00130	TIGR00229				1		1		1		3		
arCOG02333	T	Signal transduction mechanisms	-	Signal transduction histidine kinase, contains REC and PAS domains	COG00642	pfam13492,pfam0051	cd00082,cd000	TIGR02966								1				
arCOG04809	T	Signal transduction mechanisms	-	Signal transduction histidine kinase	COG00642	pfam00512,pfam0251	cd00082,cd000	TIGR02956								1				
arCOG01180	T	Signal transduction mechanisms	RIO1	Serine/threonine protein kinase involved in cell cycle control	COG01718	pfam01163	cd05145	TIGR03724	1	1		1						1	1	2
arCOG04820	T	Signal transduction mechanisms	-	SOUL heme-binding protein		pfam04832				1	1	1		1						1
arCOG06193	T	Signal transduction mechanisms	-	Signal transduction histidine kinase and PAS domains	COG00642	pfam13426,pfam1342	cd00130,cd001	TIGR00229,TIGR00229,TI		1										
arCOG04453	T	Signal transduction mechanisms	DisA_N	Diadenylate cyclase (c-di-AMP synthetase), DisA_N domain	COG01624	pfam02457												1	1	
arCOG03799	T	Signal transduction mechanisms	AtoS	GAF, PAS/PAC domains containing signal transduction protein	COG02202	pfam05763												1	2	1
arCOG00449	T	Signal transduction mechanisms	UspA	Nucleotide-binding protein, UspA family	COG00589	pfam00582,pfam0058	cd00293,cd00293											1		
arCOG01172	T	Signal transduction mechanisms	RAD55	RecA-superfamily ATPase implicated in signal transduction	COG00467	pfam06745		TIGR03881										1		
arCOG01173	T	Signal transduction mechanisms	RAD55	RecA-superfamily ATPase implicated in signal transduction	COG00467	pfam06745	cd01393	TIGR03880										1		
arCOG01174	T	Signal transduction mechanisms	RAD55	RecA-superfamily ATPase implicated in signal transduction	COG00467	pfam06745	cd01124	TIGR03877										1		
arCOG02327	T	Signal transduction mechanisms	-	Signal transduction histidine kinase, contains PAS domain	COG00642	pfam13426,pfam0251	cd00130,cd000	TIGR02966										1		
arCOG02385	T	Signal transduction mechanisms	CheY	Rec domain	COG00784	pfam00072	cd00156	TIGR01818										1		
arCOG03803	T	Signal transduction mechanisms	-	Membrane assicated inactivated KaiC-like ATPase, DUF835 family	SC.00096	pfam05763												1		
arCOG03804	T	Signal transduction mechanisms	-	Membrane assicated inactivated KaiC-like ATPase, DUF835 family	SC.00096	pfam05763												1		
arCOG04404	T	Signal transduction mechanisms	-	Cell fate regulator YlbF, YheA/YmcA/DUF963 family (controls sporulation, con	COG03679	pfam06133												1		
arCOG06538	T	Signal transduction mechanisms	-	Signal transduction histidine kinase	COG00642			TIGR00229,TIGR02956										1		
arCOG01143	T	Signal transduction mechanisms	ApaH	Serine/threonine protein phosphatase PP2A family	COG00639	pfam00149	cd00144											2	2	
arCOG01178	T	Signal transduction mechanisms	GvpD	GvpD gas vesicle protein, contains RecA/KaiC family ATPase domain	COG00467	pfam07088	cd01124	TIGR02655										2		
arCOG02391	T	Signal transduction mechanisms	CheY	Rec domain	COG00784	pfam00072	cd00156	TIGR02154										3	1	5
arCOG03413	T	Signal transduction mechanisms	CDC14	Protein-tyrosine phosphatase	COG02453	pfam00782	cd00047												1	
arCOG03517	T	Signal transduction mechanisms	-	Formylglycine-generating sulfatase enzyme	COG01262	pfam12867,pfam03781		TIGR03440											1	
arCOG00451	T	Signal transduction mechanisms	UspA	Nucleotide-binding protein, UspA family	COG00589														2	2
arCOG01992	T	Signal transduction mechanisms	SixA	Phosphohistidine phosphatase SixA	COG02062	pfam00300	cd07067	TIGR00249												1
arCOG04647	T	Signal transduction mechanisms	BolA	Stress-induced morphogen (activity unknown)	COG00271	pfam01722														1
arCOG06801	T	Signal transduction mechanisms	SrkA	Ser/Thr protein kinase RdoA involved in Cpx stress response, MazF antagonist	COG02334	pfam01636	cd05153													3
METABOLISM																				
arCOG01768	E	Amino acid transport and metabolism	GlpG	Membrane associated serine protease	COG00705	pfam01694			1	1	1	1	1		1	1	1	1	1	1
arCOG04671	E	Amino acid transport and metabolism	HutH	Histidine ammonia-lyase	COG02986	pfam00221	cd00332	TIGR01225	1	1	1	1	1		1	1	1	1	1	3
arCOG04247	E	Amino acid transport and metabolism	YpwA	Zn-dependent carboxypeptidase, M32 family	COG02317	pfam02074	cd06460		1	1	1		1	2	1	1	1	1	1	2
arCOG01352	E	Amino acid transport and metabolism	GdhA	Glutamate dehydrogenase/leucine dehydrogenase	COG00334	pfam02812,pfam0020	cd01076		1	1	1				1	1	1	1	1	2
arCOG00751	E	Amino acid transport and metabolism	DppB	ABC-type dipeptide/oligopeptide/nickel transport system, permease compon	COG00601	pfam00528	cd06261	TIGR02789	1	1		1	1		1	1	5			
arCOG03109	E	Amino acid transport and metabolism	DdaH	N-Dimethylarginine dimethylaminohydrolase	COG01834	pfam02274		TIGR01078	1	2	1	1	1	1	1	1	1	1	1	3
arCOG06678	E	Amino acid transport and metabolism	Ald	Alanine dehydrogenase	COG00686	pfam05222,pfam0126	cd05304	TIGR00561	1	2	1	1	1	1	1	1		1		5
arCOG09400	E	Amino acid transport and metabolism	PntB	NAD(P) transhydrogenase beta subunit	COG01282	pfam02233			1	2	1	1	1	1	1	1		1		5
arCOG00184	E	Amino acid transport and metabolism	AppF	ABC-type oligopeptide transport system, ATPase component	COG04608	pfam00005	cd03257	TIGR02769	1	2		1	1		1	1	4			
arCOG01700	E	Amino acid transport and metabolism	SpeB	Arginase family enzyme	COG00010	pfam00491	cd11593	TIGR01230	1		1			1	1	1	1	1	1	3
arCOG05395	E	Amino acid transport and metabolism	FtcD	Formiminotetrahydrofolate cyclodeaminase	COG03404				1		1				1	1	1	2		5
arCOG00181	E	Amino acid transport and metabolism	DppD	ABC-type dipeptide/oligopeptide/nickel transport system, ATPase component	COG00444	pfam00005	cd03257	TIGR02770	1			1	1		1	1	4			
arCOG01430	E	Amino acid transport and metabolism	CysK	Cysteine synthase	COG00031	pfam00291	cd01561	TIGR01136							1	1	1	1		
arCOG00915	E	Amino acid transport and metabolism	GabT	4-aminobutyrate aminotransferase or related aminotransferase	COG00160	pfam00202	cd00610	TIGR00707							1	1	2			
arCOG00095	E	Amino acid transport and metabolism	GltB	Glutamate synthase domain 1	COG00067	pfam00310	cd01907	TIGR01134							1	1		1		
arCOG04670	E	Amino acid transport and metabolism	HutU	Urocanate hydratase	COG02987	pfam01175		TIGR01228	1	1	1	1			1		1	1	1	2
arCOG01948	E	Amino acid transport and metabolism	RhtB	Threonine efflux protein	COG01280	pfam01810		TIGR00949	1	1	1	1			1			1	1	2
arCOG00243	E	Amino acid transport and metabolism	LYS9	Saccharopine dehydrogenase or related enzyme	COG01748	pfam03435			1	1	1	1			1			1	1	3
arCOG01158	E	Amino acid transport and metabolism	SerB	Phosphoserine phosphatase	COG00560	pfam00702		TIGR01491	1	1	1	1			1			3		2
arCOG01924	E	Amino acid transport and metabolism	AnsB	L-asparaginase/archaeal Glu-tRNAGln amidotransferase subunit D	COG00252	pfam00710	cd08962	TIGR02153	1	1		1			1		1	1	1	2
arCOG01888	E	Amino acid transport and metabolism	AmpS	Leucyl aminopeptidase (aminopeptidase T)	COG02309	pfam02073			1	1		1			1					2
arCOG04897	E	Amino acid transport and metabolism	-	Aspartate/tyrosine/aromatic aminotransferase	COG00436	pfam00155	cd00609	TIGR01141	1	2	1	1			1					
arCOG00924	E	Amino acid transport and metabolism	LivF	ABC-type branched-chain amino acid transport system, ATPase component	COG00410	pfam00005	cd03224	TIGR03410	1	2			1		1			1	1	
arCOG00925	E	Amino acid transport and metabolism	LivG	ABC-type branched-chain amino acid transport system, ATPase component	COG00411	pfam00005	cd03219	TIGR03411	1	2			1		1			1	1	
arCOG01021	E	Amino acid transport and metabolism	LivK	ABC-type branched-chain amino acid transport system, periplasmic compone	COG00683	pfam13458	cd06268		1	2			1		1			1	1	
arCOG01273	E	Amino acid transport and metabolism	LivM	ABC-type branched-chain amino acid transport system, permease component	COG04177	pfam02653	cd06581		1	2					1			1		
arCOG00912	E	Amino acid transport and metabolism	ArgF	Ornithine carbamoyltransferase	COG00078	pfam02729,pfam00185		TIGR00658	1		1	1	1	1	1		1	1	1	
arCOG04758	E	Amino acid transport and metabolism	PepF	Oligoendopeptidase F	COG01164	pfam08439,pfam0143	cd09608	TIGR00181	1		1	1	1	1	1					1
arCOG01646	E	Amino acid transport and metabolism	DAP2	Dipeptidyl aminopeptidase/acylaminoacyl-peptidase	COG01506	pfam07859	cd00312		3	2	3	2		1	1		1	2		5
arCOG00071	E	Amino acid transport and metabolism	AsnB	Asparagine synthase (glutamine-hydrolyzing)	COG00367	pfam00733	cd01991	TIGR01536		1	1	1	1	1	1		1	1	2	
arCOG00076	E	Amino acid transport and metabolism	GcvP	Glycine cleavage system protein P (pyridoxal-binding), C-terminal domain	COG01003	pfam02347	cd00613	TIGR00461		1	1	1			1			1	1	1
arCOG01319	E	Amino acid transport and metabolism	PutP	Na+/proline symporter	COG00591	pfam00474	cd11474	TIGR00813		1	1				1					1
arCOG04779	E	Amino acid transport and metabolism	IaaA	Isoaspartyl peptidase or L-asparaginase, Ntn-hydrolase superfamily	COG01446	pfam01112	cd14950			1					1		1			1
arCOG00863	E	Amino acid transport and metabolism	ArcC	Carbamate kinase	COG00549	pfam00696	cd04235	TIGR00746							1		1	1		4
arCOG01534	E	Amino acid transport and metabolism	DdpA	ABC-type transport system, periplasmic component	COG00747	pfam00496	cd08519	TIGR02294							1		3			
arCOG04333	E	Amino acid transport and metabolism	-	Aspartate/tyrosine/aromatic aminotransferase	COG00436	pfam00155	cd00609	TIGR03947							1					
arCOG00088	E	Amino acid transport and metabolism	PuuD	Gamma-glutamyl-gamma-aminobutyrate hydrolase PuuD (putrescine degrada	COG02071	pfam07722	cd01745	TIGR00888	1	1	1	1	1	1	2	1				4
arCOG00060	E	Amino acid transport and metabolism	MetC	Cystathionine beta-lyase/cystathionine gamma-synthase	COG00626	pfam01053	cd00614	TIGR02080	1	1					2	1		2	2	
arCOG01107	E	Amino acid transport and metabolism	ArgE	Acetylornithine deacetylase/Succinyl-diaminopimelate desuccinylase or relate	COG00624	pfam01546	cd05650	TIGR01910		2					2	1	2			
arCOG02706	E	Amino acid transport and metabolism	GloA	Lactoylglutathione lyase or related enzyme	COG00346	pfam13669	cd07249	TIGR03081	2	1			1	1	2	2	1	1	1	
arCOG00096	E	Amino acid transport and metabolism	GltB	Glutamate synthase domain 3	COG00070	pfam01493	cd00981	TIGR03122	1	1	1	1		1	2					1
arCOG05229	E	Amino acid transport and metabolism	PepD	Di- or tripeptidase	COG02195		cd03890	TIGR01893	1	1	1	1		1	2					
arCOG02297	E	Amino acid transport and metabolism	IlvE	Branched-chain amino acid aminotransferase/4-amino-4-deoxychorismate lya	COG00115	pfam01063	cd01558	TIGR01122	2	2	1	1			2				2	2
arCOG01000	E	Amino acid transport and metabolism	PepP	Xaa-Pro aminopeptidase	COG00006	pfam01321,pfam0055	cd01092	TIGR00500	2	1		1			3	1	3	2	5	
arCOG01130	E	Amino acid transport and metabolism	-	Aspartate/tyrosine/aromatic aminotransferase	COG00436	pfam00155	cd00609	TIGR03540	2	3	2	1			3		3	2	6	
arCOG04490	E	Amino acid transport and metabolism	PdaD	Pyruvoyl-dependent arginine decarboxylase (PvlArgDC)	COG01945	pfam01862		TIGR00286	1	1	1	1	1			1	1			
arCOG01533	E	Amino acid transport and metabolism	DdpA	ABC-type transport system, periplasmic component	COG00747	pfam00496	cd08512		1	1	1	1	2			1				
arCOG01594	E	Amino acid transport and metabolism	CarB	Carbamoylphosphate synthase large subunit	COG00458	pfam00289,pfam02786,pfam02787,pf		TIGR01369	1	1	1		1			1		1	1	
arCOG00458	E	Amino acid transport and metabolism	-	Archaemetzincin family protease	COG01913	pfam07998	cd11375		1	1		1				1				
arCOG02165	E	Amino acid transport and metabolism	-	Transglutaminase-like cysteine protease	COG01305	pfam01841			1	1						1	1	1		
arCOG00064	E	Amino acid transport and metabolism	CarA	Carbamoylphosphate synthase small subunit	COG00505	pfam00988,pfam0011	cd01744	TIGR01368	1		1		1			1		1	1	

arCOG01431	E	Amino acid transport and metabolism	IlvA	Threonine dehydratase	COG01171	pfam00291	cd01562,cd048	TIGR01127	1		1			1		1				
arCOG00923	E	Amino acid transport and metabolism	GlnQ	ABC-type polar amino acid transport system, ATPase component	COG01126		cd03262	TIGR03005						1		1				2
arCOG00771	E	Amino acid transport and metabolism	-	CubicO group peptidase, beta-lactamase class C family	COG01680	pfam00144										1				
arCOG01798	E	Amino acid transport and metabolism	HisM	ABC-type amino acid transport system, permease component	COG00765	pfam00528	cd06261	TIGR03003						2		2				3
arCOG00914	E	Amino acid transport and metabolism	ArgD	Ornithine/acetylornithine aminotransferase	COG04992	pfam00202	cd00610	TIGR00707	1	1	1	1	1						1	1
arCOG01909	E	Amino acid transport and metabolism	GlnA	Glutamine synthetase	COG00174	pfam00120		TIGR03105	1	1	1	1					2	1	2	
arCOG00756	E	Amino acid transport and metabolism	GcvT	Glycine cleavage system T protein (aminomethyltransferase)	COG00404	pfam01571,pfam08669		TIGR00528	1	1		1					1	1	1	
arCOG01303	E	Amino acid transport and metabolism	GcvH	Glycine cleavage system H protein (lipoate-binding)	COG00509	pfam01597	cd06848	TIGR00527	1	1		1					1	2	5	
arCOG06322	E	Amino acid transport and metabolism	PutA	Proline dehydrogenase	COG00506	pfam01619			1	1		1						1		
arCOG00755	E	Amino acid transport and metabolism	DadA	Glycine/D-amino acid oxidase (deaminating)	COG00665	pfam01266			1	1							1		2	
arCOG01947	E	Amino acid transport and metabolism	RhtB	Threonine efflux protein	COG01280	pfam01810			1	2										
arCOG01110	E	Amino acid transport and metabolism	ArgE	Acetylornithine deacetylase/Succinyl-diaminopimelate desuccinylase or relate	COG00624	pfam01546	cd05681	TIGR01910	1		1		1							1
arCOG02767	E	Amino acid transport and metabolism	-	Metal-dependent membrane protease, CAAX family	COG01266	pfam02517			1		1			1						
arCOG03600	E	Amino acid transport and metabolism	-	Transglutaminase-like cysteine protease	COG01305				1		1									
arCOG00177	E	Amino acid transport and metabolism	PotA	ABC-type spermidine/putrescine transport system, ATPase component	COG03842	pfam00005	cd03301	TIGR03265	1			1								1
arCOG03101	E	Amino acid transport and metabolism	GloA	Lactoylglutathione lyase or related enzyme	COG00346	pfam00903	cd08346		1			1								
arCOG14875	E	Amino acid transport and metabolism	-	Aspartyl protease		pfam05618			1				1							1
arCOG01316	E	Amino acid transport and metabolism	PutP	Na+/proline symporter	COG00591	pfam00474	cd10322	TIGR00813	2	1							1	1	3	
arCOG00027	E	Amino acid transport and metabolism	MfnA	L-tyrosine decarboxylase, PLP-dependent protein	COG00076	pfam00282	cd06450	TIGR03812		1							1	1		
arCOG07760	E	Amino acid transport and metabolism	-	Peptidase family C25		pfam01364	cd02258			1							1			
arCOG01512	E	Amino acid transport and metabolism	HyuB	N-methylhydantoinase B/acetone carboxylase, alpha subunit	COG00146	pfam02538				1										
arCOG01949	E	Amino acid transport and metabolism	ArgO	Lysine/arginine efflux permease	COG01279	pfam01810		TIGR00948		1										
arCOG05121	E	Amino acid transport and metabolism	TesA	Lysophospholipase L1 or related esterase	COG02755					1										
arCOG00121	E	Amino acid transport and metabolism	AsnB	Asparagine synthase (glutamine-hydrolyzing)	COG00367	pfam00733						1								
arCOG03999	E	Amino acid transport and metabolism	-	Clostripain family peptidase, C11		pfam03415		TIGR02806				1								
arCOG01269	E	Amino acid transport and metabolism	LivH	Branched-chain amino acid ABC-type transport system, permease component	COG00559	pfam02653	cd06582	TIGR03409					1							1
arCOG03611	E	Amino acid transport and metabolism	-	Peptidase C1A subfamily			cd08546,cd14254							1						
arCOG01596	E	Amino acid transport and metabolism	CarB	Carbamoylphosphate synthase large subunit	COG00458	pfam15632		TIGR01205						1						
arCOG00077	E	Amino acid transport and metabolism	GcvP	Glycine cleavage system protein P (pyridoxal-binding), N-terminal domain	COG00403	pfam02347	cd00613	TIGR00461									1	1	1	
arCOG01434	E	Amino acid transport and metabolism	ThrC	Threonine synthase and cysteate synthase	COG00498	pfam01155,pfam0029	cd01563	TIGR02605,TIGR00260									1	1	1	
arCOG00279	E	Amino acid transport and metabolism	SpeD	S-adenosylmethionine decarboxylase/arginine decarboxylase	COG01586	pfam02675		TIGR03330									1	1	2	
arCOG00073	E	Amino acid transport and metabolism	CysH	PAPS reductase related enzyme fused to RNA-binding PUA domain and ferred	COG00175	pfam01507	cd01713	TIGR00434									1		2	
arCOG00009	E	Amino acid transport and metabolism	PotE	Amino acid transporter	COG00531	pfam00324		TIGR00909									1			
arCOG00050	E	Amino acid transport and metabolism	SpeE	Spermidine synthase	COG00421	pfam01564	cd02440	TIGR00417									1			
arCOG00092	E	Amino acid transport and metabolism	LdcC	Arginine/lysine/ornithine decarboxylase	COG01982	pfam03709,pfam0127	cd00615	TIGR04301									1			
arCOG00304	E	Amino acid transport and metabolism	HIS2	Histidinol phosphatase or related hydrolase of the PHP family	COG01387	pfam02811	cd12110	TIGR01856									1			
arCOG01123	E	Amino acid transport and metabolism	MtnA	Methylthioribose-1-phosphate isomerase (methionine salvage pathway), a pa	COG00182	pfam01008		TIGR00512									1			
arCOG01292	E	Amino acid transport and metabolism	GltD	NADPH-dependent glutamate synthase beta chain or related oxidoreductase	COG00493	pfam07992		TIGR01316									1			
arCOG01647	E	Amino acid transport and metabolism	-	Serine protease of the peptidase family S9A	COG01505	pfam00326											1			
arCOG01678	E	Amino acid transport and metabolism	MetK	Archaeal S-adenosylmethionine synthetase	COG01812	pfam01941											1			
arCOG02167	E	Amino acid transport and metabolism	-	Transglutaminase-like cysteine protease	COG01305	pfam01841											1			
arCOG02766	E	Amino acid transport and metabolism	-	Metal-dependent membrane protease, CAAX family	COG01266	pfam02517											1			
arCOG03602	E	Amino acid transport and metabolism	PepD	Dipeptidase family C69	COG04690	pfam03577											1			
arCOG03672	E	Amino acid transport and metabolism	-	Thermopsin-like protease	SC.00065	pfam05317,pfam06832											1			
arCOG05065	E	Amino acid transport and metabolism	CdsB	L-cysteine desulfidase	COG03681	pfam03313											1			
arCOG05363	E	Amino acid transport and metabolism	-	CAAX family protease, possibly associated with type IV pili like system	COG01266	pfam02517											1			
arCOG05808	E	Amino acid transport and metabolism	DUR1	Allophanate hydrolase subunit 2	COG01984	pfam02626		TIGR00724									1			
arCOG05809	E	Amino acid transport and metabolism	DUR1	Allophanate hydrolase subunit 1	COG02049	pfam02682		TIGR00370									1			
arCOG07546	E	Amino acid transport and metabolism	-	Zincin superfamily protease													1			
arCOG10063	E	Amino acid transport and metabolism	SdaC	Tryptophan/tyrosine permease family	COG00814	pfam03222		TIGR00837									1			
arCOG13501	E	Amino acid transport and metabolism	-	Serine dehydratase beta chain		pfam03315	cd04879	TIGR00719									1			
arCOG00070	E	Amino acid transport and metabolism	GlyA	Glycine/serine hydroxymethyltransferase	COG00112	pfam00464	cd00378										2	1	1	
arCOG01698	E	Amino acid transport and metabolism	LeuC	Homoaconitate hydratase/3-isopropylmalate dehydratase large subunit famil	COG00065	pfam00330	cd01583	TIGR02086									3	1	4	
arCOG02230	E	Amino acid transport and metabolism	LeuD	3-isopropylmalate dehydratase small subunit	COG00066	pfam00694	cd01577	TIGR02087									3			
arCOG00748	E	Amino acid transport and metabolism	DppC	ABC-type dipeptide/oligopeptide/nickel transport system, permease compon	COG01173	pfam12911,pfam0052	cd06261	TIGR02790									5			
arCOG00494	E	Amino acid transport and metabolism	Asd	Aspartate-semialdehyde dehydrogenase	COG00136	pfam01118,pfam02774		TIGR00978										1	1	
arCOG02014	E	Amino acid transport and metabolism	TrpE	Anthranilate/para-aminobenzoate synthase component I	COG00147	pfam04715,pfam00425		TIGR01820										1	1	
arCOG02097	E	Amino acid transport and metabolism	AroD	3-dehydroquinate dehydratase	COG00710	pfam01487	cd00502	TIGR01093										1	1	
arCOG04134	E	Amino acid transport and metabolism	AroA	5-enolpyruvylshikimate-3-phosphate synthase	COG00128	pfam00275	cd01556	TIGR01356										1	1	
arCOG06005	E	Amino acid transport and metabolism	-	Predicted metalloprotease, contains C-terminal PDZ domain	COG03975	pfam05299,pfam1318	cd00990	TIGR02037										1	1	
arCOG07384	E	Amino acid transport and metabolism	HisB	Histidinol phosphatase or related phosphatase	COG00241	pfam13242	cd01427	TIGR01656										1	1	
arCOG01025	E	Amino acid transport and metabolism	AroK2	Archaeal shikimate kinase	COG01685	pfam00288		TIGR01920										1	2	
arCOG01131	E	Amino acid transport and metabolism	-	Aspartate/tyrosine/aromatic aminotransferase	COG00436	pfam00155	cd00609	TIGR01265										1	2	
arCOG04133	E	Amino acid transport and metabolism	AroC	Chorismate synthase	COG00082	pfam01264	cd07304	TIGR00033										1	2	
arCOG04353	E	Amino acid transport and metabolism	-	3-dehydroquinate synthase	COG01465	pfam01959												1	2	
arCOG02092	E	Amino acid transport and metabolism	LeuA	Isopropylmalate/homocitrate/citramalate synthase	COG00119	pfam00682,pfam0103	cd07940	TIGR02146										1	3	
arCOG04322	E	Amino acid transport and metabolism	PepB	Leucyl aminopeptidase	COG00260	pfam02789,pfam0088	cd00433											1	3	
arCOG01890	E	Amino acid transport and metabolism	AmpS	Leucyl aminopeptidase (aminopeptidase T)	COG02309	pfam02073												1	4	
arCOG00861	E	Amino acid transport and metabolism	LysC	Aspartokinase	COG00527													1		
arCOG01459	E	Amino acid transport and metabolism	Tdh	Threonine dehydrogenase or related Zn-dependent dehydrogenase	COG01063	pfam08240,pfam0010	cd08235	TIGR00692										1		
arCOG02769	E	Amino acid transport and metabolism	-	Metal-dependent membrane protease, CAAX family	COG01266	pfam02517												1		
arCOG04044	E	Amino acid transport and metabolism	-	2-amino-3,7-dideoxy-D-threo-hept-6-ulosonic acid synthase, DhnA-aldolase fa	COG01830	pfam01791	cd00958	TIGR01949										2	4	
arCOG00619	E	Amino acid transport and metabolism	GltB	Glutamate synthase domain 2 and ferredoxin domain	COG00069	pfam01645	cd02808												1	
arCOG01033	E	Amino acid transport and metabolism	AroE	Shikimate 5-dehydrogenase	COG00169	pfam08501,pfam0148	cd01065	TIGR00507											1	
arCOG01799	E	Amino acid transport and metabolism	HisJ	ABC-type amino acid transport/signal transduction system, periplasmic comp	COG00834	pfam00497	cd13624	TIGR01096											1	
arCOG04273	E	Amino acid transport and metabolism	HisC	Histidinol-phosphate/aromatic aminotransferase or cobyric acid decarboxylas	COG00079	pfam00155	cd00609	TIGR01140											1	
arCOG00757	E	Amino acid transport and metabolism	GcvT	Glycine cleavage system T protein (aminomethyltransferase)	COG00404	pfam01571,pfam08669		TIGR00528											2	
arCOG01122	G	Carbohydrate transport and metabolism	RpiA	Ribose 5-phosphate isomerase	COG00120	pfam06026	cd01398	TIGR00021	1	1		1			1	1	1	1	3	
arCOG00767	G	Carbohydrate transport and metabolism	{ManB}	Phosphomannomutase	COG01109	pfam02878,pfam0287	cd03087	TIGR03990	1		1	1		1	1	1	3	2	3	
arCOG00493	G	Carbohydrate transport and metabolism	GapA	Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydro	COG00057	pfam00044,pfam02800		TIGR01546	1			1			1	1	1	1	2
arCOG00052	G	Carbohydrate transport and metabolism	Pgi	Glucose-6-phosphate isomerase	COG00166	pfam01380,pfam1043	cd05017,cd056	TIGR02128	1			1			1	1		1	1	

arCOG04170	G	Carbohydrate transport and metabolism	GckA	Glycerate kinase	COG02379	pfam13660,pfam05161								1	1	1		
arCOG04961	G	Carbohydrate transport and metabolism	GlpX	Fructose-1,6-bisphosphatase/sedoheptulose 1,7-bisphosphatase or related pr	COG01494	pfam03320	cd01516	TIGR00330							1	1		1	4
arCOG00274	G	Carbohydrate transport and metabolism	RhaT	Permease of the drug/metabolite transporter (DMT) superfamily	COG00697	pfam00892,pfam00892		TIGR00950							1	1			2
arCOG01349	G	Carbohydrate transport and metabolism	SuhB	Archaeal fructose-1,6-bisphosphatase or related enzyme of inositol monopho	COG00483	pfam00459	cd01642	TIGR01331	1	1	1		1	1	1	3	1		2
arCOG00014	G	Carbohydrate transport and metabolism	RbsK	Sugar kinase, ribokinase family	COG00524	pfam00294	cd01174	TIGR02152	1	2	1	1	1	1	1	3	1	2	6
arCOG01993	G	Carbohydrate transport and metabolism	GpmA	Phosphoglycerate mutase 1	COG00588	pfam00300	cd07067	TIGR01258	1	1	1		1	2	1			1	2
arCOG04431	G	Carbohydrate transport and metabolism	GlpF	Glycerol uptake facilitator or related permease (Major Intrinsic Protein Family	COG00580	pfam00230	cd00333	TIGR00861	1	1		1			1			1	2
arCOG01087	G	Carbohydrate transport and metabolism	TpiA	Triosephosphate isomerase	COG00149	pfam00121	cd00311	TIGR00419	1	1					1		1	1	1
arCOG05046	G	Carbohydrate transport and metabolism	Rpe	Pentose-5-phosphate-3-epimerase	COG00036	pfam00834	cd00429	TIGR01163		1	1	1			1		1	1	3
arCOG00130	G	Carbohydrate transport and metabolism	-	MFS family permease	COG00477	pfam07690	cd06174			1	1	1			1		4	1	1
arCOG01169	G	Carbohydrate transport and metabolism	Eno	Enolase	COG00148	pfam03952,pfam0011	cd03313	TIGR01060		1		1	1	1	1		1	1	2
arCOG01111	G	Carbohydrate transport and metabolism	PpsA	Phosphoenolpyruvate synthase/pyruvate phosphate dikinase	COG00574	pfam01326,pfam00391,pfam02896		TIGR01418	1	2	1	1	1	1	2	1	1		
arCOG00496	G	Carbohydrate transport and metabolism	Pgk	3-phosphoglycerate kinase	COG00126	pfam00162	cd00318		1		1				2	1	1	1	1
arCOG01053	G	Carbohydrate transport and metabolism	TktA1	Transketolase, N-terminal subunit	COG03959	pfam00456	cd02012	TIGR00232	2	2	3	3		2	3	2	1	2	6
arCOG00271	G	Carbohydrate transport and metabolism	RhaT	Permease of the drug/metabolite transporter (DMT) superfamily	COG00697	pfam00892,pfam00892		TIGR00950	4	6	1	1	1	1	3	2	2	3	3
arCOG01420	G	Carbohydrate transport and metabolism	GlgA	Glycogen synthase	COG00297	pfam08323,pfam0053	cd03791	TIGR02095	1	1		1				1			
arCOG03460	G	Carbohydrate transport and metabolism	LmbE	N-acetylglucosaminyl deacetylase, LmbE family	COG02120	pfam02585		TIGR04001	1	2						1	1		
arCOG07581	G	Carbohydrate transport and metabolism	-	Glycosyl hydrolase family 18, contains cellulose binding domain	COG03325	pfam02839,pfam0080	cd12215,cd00146,cd06548		3	3	1	2				1	2	4	4
arCOG02796	G	Carbohydrate transport and metabolism	-	Glucose/sorbosone dehydrogenase	COG02133	pfam07995		TIGR03606								1		1	1
arCOG00138	G	Carbohydrate transport and metabolism	-	MFS family permease	COG00477				1	1	1	1	1						
arCOG03641	G	Carbohydrate transport and metabolism	PfkA	6-phosphofructokinase	COG00205	pfam00365	cd00363	TIGR02483	1	1	1	1		1					
arCOG02876	G	Carbohydrate transport and metabolism	CDA1	Peptidoglycan/xylan/chitin deacetylase, PgdA/CDA1 family	COG00726	pfam01522	cd10941	TIGR03006	1		1								
arCOG00147	G	Carbohydrate transport and metabolism	-	MFS family permease	COG00477	pfam07690	cd06174	TIGR00890	1			1		1					2
arCOG04934	G	Carbohydrate transport and metabolism	-	Peptidoglycan/xylan/chitin deacetylase, PgdA/CDA1 family	COG00726				1										
arCOG00143	G	Carbohydrate transport and metabolism	-	MFS family permease	COG00477	pfam07690	cd06174	TIGR00711		1								1	1
arCOG07840	G	Carbohydrate transport and metabolism	-	Glycosyl hydrolase family 18, contains cellulose binding domain	COG03325	pfam02839,pfam0080	cd12215,cd00146,cd06543			1									2
arCOG00754	G	Carbohydrate transport and metabolism	LhgO	L-2-hydroxyglutarate oxidase LhgO	COG00579	pfam01266,pfam04324		TIGR03377						1			1		
arCOG04221	G	Carbohydrate transport and metabolism	NagD	Phosphatase of the HAD superfamily	COG00647	pfam13344,pfam1324	cd01427	TIGR01457									1	1	2
arCOG00135	G	Carbohydrate transport and metabolism	-	MFS family permease	COG00477	pfam05977	cd06174	TIGR00900									1	1	3
arCOG01051	G	Carbohydrate transport and metabolism	TktA2	Transketolase, C-terminal subunit	COG03958	pfam02779,pfam0278	cd07033	TIGR00204									1	1	3
arCOG01029	G	Carbohydrate transport and metabolism	GalK	Galactokinase	COG00153	pfam10509,pfam08544		TIGR00131									1	1	
arCOG01393	G	Carbohydrate transport and metabolism	-	Glycosyl transferase, related to UDP-glucuronosyltransferase	COG01819												1	1	
arCOG04120	G	Carbohydrate transport and metabolism	PykF	Pyruvate kinase	COG00469	pfam00224,pfam0288	cd00288	TIGR01064									1	1	
arCOG05061	G	Carbohydrate transport and metabolism	TalA	Transaldolase	COG00176	pfam00923	cd00956	TIGR00875									1	1	
arCOG05412	G	Carbohydrate transport and metabolism	BglB	Beta-glucosidase/6-phospho-beta-glucosidase/beta-galactosidase	COG02723	pfam00232		TIGR03356									1	1	
arCOG00132	G	Carbohydrate transport and metabolism	-	MFS family permease	COG00477	pfam07690	cd06174										1		
arCOG00136	G	Carbohydrate transport and metabolism	-	MFS family permease	COG00477	pfam13347	cd06174	TIGR01301,TIGR00895									1		
arCOG00142	G	Carbohydrate transport and metabolism	-	MFS family permease	COG00477	pfam07690	cd06174	TIGR00711									1		
arCOG00144	G	Carbohydrate transport and metabolism	-	MFS family permease	COG00477	pfam07690	cd06174	TIGR00711									1		
arCOG00272	G	Carbohydrate transport and metabolism	RhaT	Permease of the drug/metabolite transporter (DMT) superfamily	COG00697	pfam00892											1		
arCOG00760	G	Carbohydrate transport and metabolism	Mcl/CitE	Beta-methylmalyl-CoA lyase, Citrate lyase beta subunit family	COG02301	pfam03328		TIGR01588									1		
arCOG01696	G	Carbohydrate transport and metabolism	ApgM	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	COG03635	pfam01676		TIGR00306									1		
arCOG02602	G	Carbohydrate transport and metabolism	OxdD	Oxalate decarboxylase/archaeal phosphoglucose isomerase, cupin superfamil	COG02140	pfam06560											1		
arCOG02682	G	Carbohydrate transport and metabolism	-	MFS family permease	COG00477	pfam00083	cd06174	TIGR00887									1		
arCOG03278	G	Carbohydrate transport and metabolism	-	Glycosyl hydrolase family 57	COG01449	pfam03065	cd10795										1		
arCOG04180	G	Carbohydrate transport and metabolism	-	Bifunctional fructose-1,6-bisphosphate aldolase/phosphatase FBPA/FBPase	COG01980	pfam01950											1		
arCOG04226	G	Carbohydrate transport and metabolism	AraD	Ribulose-5-phosphate 4-epimerase/Fuculose-1-phosphate aldolase	COG00235	pfam00596	cd00398	TIGR03328									1		
arCOG04443	G	Carbohydrate transport and metabolism	RbcL	Ribulose 1,5-bisphosphate carboxylase, large subunit	COG01850	pfam02788,pfam0001	cd08213	TIGR03326									1		
arCOG05728	G	Carbohydrate transport and metabolism	-	MFS family permease	COG00477	pfam07690	cd06174	TIGR00881									1		
arCOG00139	G	Carbohydrate transport and metabolism	-	MFS family permease	COG00477	pfam07690	cd06174										2		
arCOG01518	G	Carbohydrate transport and metabolism	FrvX	Cellulase M or related protein	COG01363	pfam05343	cd05656	TIGR03107									3		
arCOG00210	G	Carbohydrate transport and metabolism	TagH	ABC-type polysaccharide/polyol phosphate transport system, ATPase compon	COG01134	pfam00005	cd03220	TIGR01188										1	1
arCOG04339	G	Carbohydrate transport and metabolism	TagG	ABC-type polysaccharide/polyol phosphate export systems, permease compo	COG01682	pfam01061												1	1
arCOG01967	G	Carbohydrate transport and metabolism	Pgk2	2-phosphoglycerate kinase	COG02074													1	2
arCOG00175	G	Carbohydrate transport and metabolism	MalK	ABC-type sugar transport system, ATPase component	COG03839	pfam00005	cd03299	TIGR03265										1	
arCOG05355	G	Carbohydrate transport and metabolism	GmhA	Phosphoheptose isomerase	COG00279														1
arCOG01490	H	Coenzyme transport and metabolism	FolA	Dihydrofolate reductase	COG00262	pfam00186	cd00209		1	1		1	1		1	1		1	4
arCOG01484	H	Coenzyme transport and metabolism	RibD	Pyrimidine reductase, riboflavin biosynthesis	COG01985	pfam01872		TIGR01508	1	1		1			1	1	1		
arCOG04262	H	Coenzyme transport and metabolism	-	Phosphopantothenate synthetase	COG01701	pfam02006			1	1					1	1	1	1	1
arCOG04538	H	Coenzyme transport and metabolism	FolD	5,10-methylene-tetrahydrofolate dehydrogenase/Methenyl tetrahydrofolate	COG00190	pfam00763,pfam0288	cd01080		1	2	1	1		1	1	1	1	1	1
arCOG00476	H	Coenzyme transport and metabolism	UbiA	4-hydroxybenzoate polyprenyltransferase or related prenyltransferase	COG00382	pfam01040	cd13961	TIGR01476	1	3		1			1	1	1	1	
arCOG00034	H	Coenzyme transport and metabolism	PDX2	Predicted glutamine amidotransferase involved in pyridoxine biosynthesis	COG00311	pfam01174	cd01749	TIGR03800	1		1		1		1	1	2	1	2
arCOG00972	H	Coenzyme transport and metabolism	NadR	Nicotinamide mononucleotide adenylyltransferase	COG01056	pfam01467	cd02166	TIGR01527	1			1		1	1	1	1	1	1
arCOG01223	H	Coenzyme transport and metabolism	CAB4	Phosphopantetheine adenylyltransferase	COG01019	pfam01467	cd02164	TIGR00125	1			1			1	1	1		1
arCOG01940	H	Coenzyme transport and metabolism	BirA	Biotin-(acetyl-CoA carboxylase) ligase	COG00340	pfam03099		TIGR00121	1			1			1	1	1		2
arCOG00584	H	Coenzyme transport and metabolism	PanB	Ketopantoate hydroxymethyltransferase	COG00413	pfam02548	cd06557	TIGR00222	2	1	1	1			1	1		1	2
arCOG04139	H	Coenzyme transport and metabolism	ApbA	Ketopantoate reductase	COG01893	pfam02558,pfam08546		TIGR00745			1				1	1	1	1	2
arCOG01320	H	Coenzyme transport and metabolism	RibB	3,4-dihydroxy-2-butanone 4-phosphate synthase	COG00108	pfam00926		TIGR00506	1	1	1	1		1	1		1	1	3
arCOG01323	H	Coenzyme transport and metabolism	RibH	Riboflavin synthase beta-chain	COG00054	pfam00885	cd09211	TIGR00114	1	1	1	1		1	1		1	1	3
arCOG01671	H	Coenzyme transport and metabolism	UbiD	3-polyprenyl-4-hydroxybenzoate decarboxylase or related decarboxylase	COG00043	pfam01977		TIGR03701	1	1	1	1		1	1		1	2	3
arCOG04713	H	Coenzyme transport and metabolism	RibC	Riboflavin synthase alpha chain	COG00307	pfam00677,pfam00677		TIGR00187	1	1	1	1		1	1			1	3
arCOG01703	H	Coenzyme transport and metabolism	UbiX	3-polyprenyl-4-hydroxybenzoate decarboxylase	COG00163	pfam02441		TIGR00421	1	1	1			1	1		1	1	1
arCOG00040	H	Coenzyme transport and metabolism	Hpt1	Hypoxanthine phosphoribosyltransferase	COG02236	pfam00156	cd06223	TIGR01251	1	1		1			1		1		
arCOG00478	H	Coenzyme transport and metabolism	UbiA	4-hydroxybenzoate polyprenyltransferase or related prenyltransferase	COG00382	pfam01040	cd13956	TIGR01475	1	2	1	1			1				
arCOG00113	H	Coenzyme transport and metabolism	BioF	7-keto-8-aminopelargonate synthetase or related enzyme	COG00156	pfam00155	cd06454	TIGR01825	1	2					1		1	1	2
arCOG02939	H	Coenzyme transport and metabolism	PhhB	Pterin-4a-carbinolamine dehydratase	COG02154	pfam01329	cd00488		2	2	1	2	2	1	1				2
arCOG01321	H	Coenzyme transport and metabolism	RibA	GTP cyclohydrolase II	COG00807	pfam00926,pfam0092	cd00641	TIGR00506,TIGR00505		1	1				1			2	4
arCOG01482	H	Coenzyme transport and metabolism	NadC	Nicotinate-nucleotide pyrophosphorylase	COG00157	pfam02749,pfam0172	cd01572	TIGR00078		1		1	1	1	1		1	1	2
arCOG04459	H	Coenzyme transport and metabolism	NadA	Quinolinate synthase	COG00379	pfam02445		TIGR00550		1					1		1	1	4
arCOG04137	H	Coenzyme transport and metabolism	SAM1	S-adenosylhomocysteine hydrolase	COG00499	pfam05221	cd00401	TIGR00936							1		1	1	2
arCOG07444	H	Coenzyme transport and metabolism	MetK	S-adenosylmethionine synthetase	COG00192	pfam00438,pfam02772,pfam02773		TIGR01034							1			1	1

arCOG03838	H	Coenzyme transport and metabolism	PqqD	Coenzyme PQQ synthesis protein D		pfam05402			1	1	1	1	1	1	2	1		1	
arCOG00069	H	Coenzyme transport and metabolism	NadE	NH3-dependent NAD+ synthetase	COG00171	pfam02540	cd00553	TIGR00552	1	2	1	1	1	1	2	1	1	1	3
arCOG01726	H	Coenzyme transport and metabolism	IspA	Geranylgeranyl pyrophosphate synthase	COG00142	pfam00348	cd00685	TIGR02748	2	2	2	2	1	1	2	1	1	2	5
arCOG01942	H	Coenzyme transport and metabolism	LipB	Lipoate-protein ligase B	COG00321	pfam03099		TIGR00214	1	1	1	1	1		2			1	2
arCOG00660	H	Coenzyme transport and metabolism	LipA	Lipoate synthase	COG00320	pfam04055	cd01335	TIGR00510	1	1		1			2			1	2
arCOG01904	H	Coenzyme transport and metabolism	-	CTP-dependent Riboflavin kinase	COG01339	pfam13412,pfam14277,pfam01982			2	1	1		1		2		1	1	2
arCOG00480	H	Coenzyme transport and metabolism	MenA	1,4-dihydroxy-2-naphthoate octaprenyltransferase	COG01575	pfam01040	cd13962	TIGR02235	1	1	1	1	1			1			
arCOG00489	H	Coenzyme transport and metabolism	PduO	Cob(I)alamin adenosyltransferase	COG02096	pfam01923		TIGR00636	1	1	1			1		1		1	
arCOG02172	H	Coenzyme transport and metabolism	QueD	6-pyruvoyl-tetrahydropterin synthase	COG00720	pfam01242	cd00470	TIGR03367	1	1		1				1	1	1	1
arCOG01045	H	Coenzyme transport and metabolism	CoaE	Dephospho-CoA kinase	COG00237		cd02022		1	1		1				1	1	1	2
arCOG02817	H	Coenzyme transport and metabolism	FolC	Folylpolyglutamate synthase and Dihydropteroate synthase	COG00285	pfam08245,pfam0287	cd00739	TIGR01499,TIGR01496		1	1	1		1		1	2	1	1
arCOG01704	H	Coenzyme transport and metabolism	Dfp	Phosphopantothenoylcysteine synthetase/decarboxylase	COG00452	pfam02441,pfam04127		TIGR00521		1						1	1	1	2
arCOG00214	H	Coenzyme transport and metabolism	MoaB	Molybdopterin biosynthesis enzyme	COG00521	pfam00994	cd00886	TIGR02667								1	1	1	2
arCOG00534	H	Coenzyme transport and metabolism	MoaE	Molybdopterin converting factor, large subunit	COG00314	pfam02391	cd00756									1	1	1	4
arCOG00930	H	Coenzyme transport and metabolism	MoaA	Molybdenum cofactor biosynthesis enzyme	COG02896	pfam04055,pfam0646	cd01335	TIGR02668								1	1	1	4
arCOG01530	H	Coenzyme transport and metabolism	MoaC	Molybdenum cofactor biosynthesis enzyme	COG00315	pfam01967	cd01419	TIGR00581								1	1	1	4
arCOG01872	H	Coenzyme transport and metabolism	MobA	Molybdopterin-guanine dinucleotide biosynthesis protein A	COG00746	pfam12804	cd02503									1	1		2
arCOG00217	H	Coenzyme transport and metabolism	MoeA	Molybdopterin biosynthesis enzyme	COG00303	pfam03453,pfam0099	cd00887	TIGR00177								1	1		4
arCOG02199	H	Coenzyme transport and metabolism	Mcr1	NAD(P)H-flavin reductase	COG00543											1	3		
arCOG04263	H	Coenzyme transport and metabolism	-	Pantoate kinase	COG01829	pfam00288			1	1	1	1					1	1	2
arCOG00021	H	Coenzyme transport and metabolism	-	Predicted transcriptional regulator fused phosphomethylpyrimidine kinase, in	COG01992	pfam10120			1	1	1	1						1	
arCOG04075	H	Coenzyme transport and metabolism	SNZ1	Pyridoxine biosynthesis enzyme	COG00214	pfam01680,pfam0569	cd04727	TIGR00343	1	2							1	1	3
arCOG00226	H	Coenzyme transport and metabolism	TbpA	ABC-type thiamine transport system, periplasmic component	COG04143	pfam13343	cd13545	TIGR01254	1			1						1	
arCOG01589	H	Coenzyme transport and metabolism	RimK	Glutathione synthase/glutaminyl transferase/alpha-L-glutamate ligase	COG00189	pfam08443		TIGR02144	1				1					1	2
arCOG04813	H	Coenzyme transport and metabolism	PanD	Aspartate 1-decarboxylase	COG00853	pfam02261	cd06919	TIGR00223	1										1
arCOG03402	H	Coenzyme transport and metabolism	MtbC1	Methanogenic corrinoid protein MtbC1	COG05012	pfam02607,pfam0231	cd02065	TIGR02370	1										
arCOG04542	H	Coenzyme transport and metabolism	FolE	GTP cyclohydrolase I	COG00302	pfam01227	cd00642	TIGR00063		1								1	1
arCOG01522	H	Coenzyme transport and metabolism	HemY	Protoporphyrinogen oxidase	COG01232	pfam01593		TIGR00562		1									
arCOG01676	H	Coenzyme transport and metabolism	ThiF	Dinucleotide-utilizing enzyme involved in molybdopterin and thiamine biosyn	COG00476	pfam00899,pfam0523	cd00757	TIGR02356									1	1	2
arCOG01348	H	Coenzyme transport and metabolism	nadF	NAD kinase	COG00061	pfam01513											1	1	
arCOG01754	H	Coenzyme transport and metabolism	SerA	Phosphoglycerate dehydrogenase or related dehydrogenase	COG00111	pfam00389	cd05303	TIGR01327									1	1	
arCOG00020	H	Coenzyme transport and metabolism	ThiD	Hydroxymethylpyrimidine/phosphomethylpyrimidine kinase	COG00351	pfam08543	cd01169	TIGR00097									1		
arCOG00117	H	Coenzyme transport and metabolism	MenG	Demethylmenaquinone methyltransferase	COG00684	pfam00215,pfam0373	cd04726	TIGR03128,TIGR01935									1		
arCOG00228	H	Coenzyme transport and metabolism	MopI	Molybdopterin-binding protein	COG03585	pfam03459		TIGR00638									1		
arCOG00532	H	Coenzyme transport and metabolism	MobB	Molybdopterin-guanine dinucleotide biosynthesis protein	COG01763	pfam03205	cd03116	TIGR00176									1		
arCOG00574	H	Coenzyme transport and metabolism	THI4	Archaeal ribulose 1,5-bisphosphate synthetase/yeast thiazole synthase	COG01635	pfam01946		TIGR00292									1		
arCOG00613	H	Coenzyme transport and metabolism	LldD	FMN-dependent dehydrogenase, includes L-lactate dehydrogenase and type I	COG01304	pfam01070	cd02811	TIGR02151									1		
arCOG00638	H	Coenzyme transport and metabolism	ThiL	Thiamine monophosphate kinase	COG00611	pfam00586	cd02194	TIGR01379									1		
arCOG01089	H	Coenzyme transport and metabolism	ThiE	Thiamine monophosphate synthase	COG00352	pfam02581	cd00564	TIGR00693									1		
arCOG01322	H	Coenzyme transport and metabolism	RibC2	Archaeal riboflavin synthase	COG01731	pfam00885	cd09210	TIGR01506									1		
arCOG01481	H	Coenzyme transport and metabolism	PncB	Nicotinic acid phosphoribosyltransferase	COG01488	pfam02749,pfam0172	cd01571	TIGR01513									1		
arCOG01939	H	Coenzyme transport and metabolism	LplA	Lipoate-protein ligase A	COG00095	pfam03099		TIGR00545									1		
arCOG02741	H	Coenzyme transport and metabolism	ThiC	Thiamine biosynthesis protein ThiC	COG00422	pfam01964		TIGR00190									1		
arCOG03837	H	Coenzyme transport and metabolism	-	Lipoate-protein ligase A associated domain	COG00095												1		
arCOG04301	H	Coenzyme transport and metabolism	MptA	Fe(2+)-dependent GTP cyclohydrolase	COG01469	pfam02649		TIGR00294									1		
arCOG04536	H	Coenzyme transport and metabolism	ArfB	Creatinine amidohydrolase/Fe(II)-dependent formamide hydrolase involved in	COG01402	pfam02633		TIGR03964									1		
arCOG04678	H	Coenzyme transport and metabolism	BtuR	ATP:corrinoid adenosyltransferase	COG02109	pfam02572	cd00561	TIGR00708									1		
arCOG00536	H	Coenzyme transport and metabolism	MoaD	Molybdopterin converting factor, small subunit	COG01977	pfam02391	cd00754,cd00756										2		
arCOG02250	H	Coenzyme transport and metabolism	EcfT	Energy-coupling factor transporter transmembrane protein EcfT	COG00619												2		
arCOG01485	H	Coenzyme transport and metabolism	RibD	Pyrimidine deaminase and reductase	COG00117	pfam00383,pfam0187	cd01284	TIGR00326										1	1
arCOG01873	H	Coenzyme transport and metabolism	MobA	GT-A family glycosyltransferase involved in molybdopterin guanine dinucleoti	COG02068	pfam12804	cd04182	TIGR03310										1	1
arCOG04348	H	Coenzyme transport and metabolism	UbiE	Ubiquinone/menaquinone biosynthesis C-methylase UbiE	COG02226	pfam08241		TIGR01934										1	1
arCOG00477	H	Coenzyme transport and metabolism	UbiA	4-hydroxybenzoate polyprenyltransferase or related prenyltransferase	COG00382	pfam01040	cd13959	TIGR01475										1	2
arCOG00572	H	Coenzyme transport and metabolism	NadB	Aspartate oxidase	COG00029	pfam00890,pfam02910		TIGR00551										1	4
arCOG02624	H	Coenzyme transport and metabolism	PaaK	Coenzyme F390 synthetase	COG01541			TIGR02304										1	
arCOG00656	H	Coenzyme transport and metabolism	ThiH	Thiamine biosynthesis enzyme ThiH, FO synthase or related uncharacterized	COG01060	pfam04055		TIGR00423										2	7
arCOG04299	H	Coenzyme transport and metabolism	HemC	Porphobilinogen deaminase	COG00181	pfam01379,pfam0390	cd13644	TIGR00212											1
arCOG01548	C	Energy production and conversion	NuoD	NADH dehydrogenase subunit D	COG00649	pfam00346		TIGR01962	1	1	1	1			1	1	1	1	3
arCOG01551	C	Energy production and conversion	NuoC	NADH dehydrogenase subunit C	COG00852	pfam00329		TIGR01961	1	1	1	1			1	1	1	1	3
arCOG01554	C	Energy production and conversion	NuoB	F420H2 dehydrogenase subunit, related to NADH:ubiquinone oxidoreductase	COG00377	pfam01058		TIGR01957	1	1	1	1			1	1	1	1	3
arCOG01543	C	Energy production and conversion	NuoI	NADH dehydrogenase subunit I	COG01143	pfam12838		TIGR01971	1	1	1	1			1	1	2	1	2
arCOG01538	C	Energy production and conversion	NuoM	NADH dehydrogenase subunit M	COG01008	pfam00361		TIGR01972	1	1	1	1			1	1		1	2
arCOG01539	C	Energy production and conversion	NuoL	NADH dehydrogenase subunit L	COG01009	pfam00662,pfam00361		TIGR01974	1	1	1	1			1	1		1	2
arCOG01540	C	Energy production and conversion	NuoN	NADH dehydrogenase subunit N	COG01007	pfam00361		TIGR01770	1	1	1	1			1	1		1	2
arCOG03073	C	Energy production and conversion	NuoK	NADH dehydrogenase subunit 4L (K,kappa)	COG00713				1	1	1	1			1	1		1	2
arCOG04654	C	Energy production and conversion	NuoJ	NADH dehydrogenase subunit J	COG00839				1	1	1	1			1	1		1	2
arCOG01546	C	Energy production and conversion	NuoH	NADH dehydrogenase subunit H	COG01005	pfam00146			1	1	1	1			1	1		1	3
arCOG01557	C	Energy production and conversion	NuoA	NADH dehydrogenase subunit A	COG00838				1	1	1	1			1	1		1	3
arCOG02095	C	Energy production and conversion	OadA1	Pyruvate/oxaloacetate carboxyltransferase	COG05016	pfam00682,pfam0243	cd07937	TIGR01108	1	1		1	1	1	1	1	2	2	3
arCOG09401	C	Energy production and conversion	-	NAD(P) transhydrogenase, subunit alpha		pfam12769		TIGR00561	1	2	1	1	1	1	1	1		1	5
arCOG01697	C	Energy production and conversion	AcnA	Aconitase A	COG01048	pfam00330,pfam0069	cd01586,cd015	TIGR01341	1	2	1	1	1	2	1	1		1	1
arCOG06720	C	Energy production and conversion	FixC	Dehydrogenase (flavoprotein)	COG00644				1		1	1		1	1	1		1	4
arCOG04138	C	Energy production and conversion	NtpI	Archaeal/vacuolar-type H+-ATPase subunit I	COG01269	pfam01496			1		1	1			1	1	1	1	2
arCOG01721	C	Energy production and conversion	QcrB	Cytochrome b subunit of the bc complex	COG01290	pfam13631	cd00284		1		1		1	2	1	1		1	2
arCOG00869	C	Energy production and conversion	NtpE	Archaeal/vacuolar-type H+-ATPase subunit E	COG01390				1			1			1	1	1	1	1
arCOG02455	C	Energy production and conversion	AtpK	Archaeal/vacuolar-type Na+/H+-ATPase, subunit K	COG00636				1			1			1	1	1	1	1
arCOG02459	C	Energy production and conversion	NtpC	Archaeal/vacuolar-type H+-ATPase subunit C	COG01527				1			1			1	1	1	1	1
arCOG04102	C	Energy production and conversion	NtpG	Archaeal/vacuolar-type H+-ATPase subunit F	COG01436	pfam01990			1			1			1	1	1	1	1
arCOG00865	C	Energy production and conversion	NtpB	Archaeal/vacuolar-type H+-ATPase subunit B	COG01156	pfam02874,pfam0000	cd01135	TIGR01041	1						1	1	1	1	1
arCOG00868	C	Energy production and conversion	NtpA	Archaeal/vacuolar-type H+-ATPase subunit A	COG01155	pfam02874,pfam0000	cd01134	TIGR01043	1						1	1	1	1	1
arCOG02200	C	Energy production and conversion	Fpr	Flavodoxin reductase (ferredoxin-NADPH reductase) family 1	COG01018		cd00322	TIGR01941			1				1	1		1	

arCOG00982	C	Energy production and conversion	GldA	Glycerol dehydrogenase or related enzyme	COG00371	pfam13685	cd08173	TIGR01357							1	1	1		
arCOG01755	C	Energy production and conversion	LdhA	Lactate dehydrogenase or related 2-hydroxyacid dehydrogenase	COG01052	pfam00389	cd05301	TIGR01327							1	1	2		
arCOG02460	C	Energy production and conversion	NapF	Ferredoxin	COG01145	pfam09383,pfam1283	cd07030	TIGR02314,TIGR04041							1	1			
arCOG01749	C	Energy production and conversion	FumC	Fumarase	COG00114	pfam00206,pfam1041	cd01596	TIGR00979	1	1				1	1	2		1	1
arCOG01337	C	Energy production and conversion	SucC	Succinyl-CoA synthetase, beta subunit	COG00045	pfam08442,pfam00549		TIGR01016	1	1	1	1			1		1	1	1
arCOG01339	C	Energy production and conversion	SucD	Succinyl-CoA synthetase, alpha subunit	COG00074	pfam02629,pfam00549		TIGR01019	1	1	1	1			1		1	1	1
arCOG01052	C	Energy production and conversion	AcoB	Pyruvate/2-oxoglutarate/acetoin dehydrogenase complex, dehydrogenase (E1	COG00022	pfam02779,pfam0278	cd07036	TIGR00204	1	1	1	1			1			1	2
arCOG05014	C	Energy production and conversion	HdrE	CoB--CoM heterodisulfide reductase subunit E	COG02181				1	1	1		1		1			1	2
arCOG01054	C	Energy production and conversion	AcoA	Pyruvate/2-oxoglutarate/acetoin dehydrogenase complex, dehydrogenase (E1	COG01071	pfam00676	cd02000	TIGR03181	1	1		1			1			2	5
arCOG06073	C	Energy production and conversion	PckA	Phosphoenolpyruvate carboxykinase (ATP)	COG01866	pfam01293	cd00484	TIGR00224	1	2	1				1			1	1
arCOG00246	C	Energy production and conversion	Mdh	Malate/lactate dehydrogenase	COG00039	pfam00056,pfam0286	cd00300	TIGR01771	1	2	1				1			1	5
arCOG02921	C	Energy production and conversion	PetE	Plastocyanin	COG03794	pfam00127	cd04220	TIGR03102	1	3		1	1	2	1			1	2
arCOG01606	C	Energy production and conversion	PorA	Pyruvate:ferredoxin oxidoreductase or related 2-oxoacid:ferredoxin oxidored	COG00674	pfam01558,pfam0185	cd07034	TIGR03710	1					1	1			1	1
arCOG04101	C	Energy production and conversion	NtpD	Archaeal/vacuolar-type H+-ATPase subunit D	COG01394	pfam01813		TIGR00309	1						1		1	1	1
arCOG04949	C	Energy production and conversion	OVP1	Na+ or H+-translocating membrane pyrophosphatase	COG03808	pfam03030		TIGR01104	1						1		1	1	3
arCOG01706	C	Energy production and conversion	AceF	Pyruvate/2-oxoglutarate dehydrogenase complex, dihydrolipoamide acyltrans	COG00508	pfam00198		TIGR01349	2	1	1	2			1			1	3
arCOG01068	C	Energy production and conversion	Lpd	Pyruvate/2-oxoglutarate dehydrogenase complex, dihydrolipoamide dehydro	COG01249	pfam07992,pfam02852		TIGR01350		1	1		1		1		2	2	3
arCOG02077	C	Energy production and conversion	IscU	NifU homolog involved in Fe-S cluster formation	COG00822	pfam01592	cd06664	TIGR03419		1	1				1				
arCOG04279	C	Energy production and conversion	-	Swiveling domain associated with predicted aconitase	COG01786	pfam01989	cd01356								1		1		
arCOG04073	C	Energy production and conversion	-	MinD superfamily P-loop ATPase containing an inserted ferredoxin domain	COG01149	pfam01656	cd03110	TIGR01969							1		2		
arCOG00573	C	Energy production and conversion	SdhA	Succinate dehydrogenase/fumarate reductase, flavoprotein subunit	COG01053	pfam00890,pfam02910		TIGR02061							1				
arCOG02179	C	Energy production and conversion	NapF	Polyferredoxin	COG01145	pfam12838,pfam12838,pfam12838		TIGR01971,TIGR04105							1				
arCOG02618	C	Energy production and conversion	-	Ferredoxin	COG01146	pfam13187,pfam12139		TIGR02060							1				
arCOG00447	C	Energy production and conversion	FixB	Electron transfer flavoprotein, alpha subunit	COG02025	pfam01012,pfam0076	cd01985		1	1	1		1		2	1		1	2
arCOG01235	C	Energy production and conversion	CyoA	Heme/copper-type cytochrome/quinol oxidase, subunit 2	COG01622	pfam00116	cd13918	TIGR02866	1	1	1		1		2	1			1
arCOG02017	C	Energy production and conversion	RutF	NADH-FMN oxidoreductase RutF, flavin reductase (DIM6/NTAB) family	COG01853	pfam01613		TIGR03615	2	1	1				2	1	1	1	1
arCOG04548	C	Energy production and conversion	-	Ferredoxin	COG01146	pfam13237		TIGR02512		1					2	1			
arCOG00446	C	Energy production and conversion	FixA	Electron transfer flavoprotein, beta subunit	COG02086	pfam01012	cd01714		1	1	1		1		2			1	2
arCOG04237	C	Energy production and conversion	GltA	Citrate synthase	COG00372	pfam00285	cd06118	TIGR01800	1	2	1	1	1	1	2		1	1	2
arCOG01237	C	Energy production and conversion	CyoB	Heme/copper-type cytochrome/quinol oxidase, subunit 1 and 3	COG00843	pfam00115,pfam0051	cd01662,cd003	TIGR02891,TIGR	1	1	1		1		3	2		1	1
arCOG00332	C	Energy production and conversion	GlpC	Membrane associated Fe-S oxidoreductase	COG00247	pfam02754									3				
arCOG01252	C	Energy production and conversion	PutA	Lactaldehyde dehydrogenase, Succinate semialdehyde dehydrogenase or oth	COG01012	pfam00171	cd07088	TIGR01780	3	4	3	2		2	5	1		4	11
arCOG00570	C	Energy production and conversion	FixC	Dehydrogenase (flavoprotein)	COG00644	pfam13450		TIGR02032	8	9	4	5	3	5	8	4	2	4	9
arCOG04594	C	Energy production and conversion	QcrB	Cytochrome b subunit of the bc complex	COG01290				1		1		1	2		1			
arCOG00288	C	Energy production and conversion	NfnB	Nitroreductase	COG00778	pfam00881	cd02150	TIGR02476		2						1	1	1	2
arCOG02929	C	Energy production and conversion	PetE	Plastocyanin	COG03794	pfam00127	cd13921	TIGR02657				1				1			
arCOG00333	C	Energy production and conversion	GlpC	Fe-S oxidoreductase	COG00247	pfam13534,pfam02754										1	1	1	2
arCOG02187	C	Energy production and conversion	NapF	Ferredoxin domain containing protein	COG01145	pfam12838,pfam12838		TIGR04105,TIGR04041								1			
arCOG02398	C	Energy production and conversion	CcdA	Cytochrome c biogenesis protein	COG00785	pfam02683										1			
arCOG02449	C	Energy production and conversion	NapF	Flavodoxin fused to ferredoxin domain	COG01145	pfam12724,pfam13187		TIGR04041								1			
arCOG01458	C	Energy production and conversion	Qor	NADPH:quinone reductase or related Zn-dependent oxidoreductase	COG00604	pfam08240,pfam0010	cd08264	TIGR02824	1	2	2		2	1		2		2	2
arCOG01491	C	Energy production and conversion	BisC	Anaerobic dehydrogenase	COG00243	pfam00384	cd02766	TIGR01591								3	1	1	3
arCOG01500	C	Energy production and conversion	HybA	Fe-S-cluster-containing dehydrogenase component	COG00437	pfam13459,pfam13247		TIGR02951								3			
arCOG00963	C	Energy production and conversion	GlpC	Succinate dehydrogenase/fumarate reductase, Fe-S protein subunit	COG00479	pfam13085,pfam13183,pfam02754,pf		TIGR00384,TIGR	1	1	1	1						1	2
arCOG01164	C	Energy production and conversion	Icd	Isocitrate dehydrogenase	COG00538	pfam00180		TIGR00183	1	1		1						1	
arCOG04650	C	Energy production and conversion	CyoC	Heme/copper-type cytochrome/quinol oxidase, subunit 3	COG01845	pfam00510	cd00386	TIGR02842	1	1			1					1	1
arCOG01599	C	Energy production and conversion	PorB	Pyruvate:ferredoxin oxidoreductase or related 2-oxoacid:ferredoxin oxidored	COG01013	pfam02775,pfam1236	cd03375	TIGR02177	1	1				1			1	1	1
arCOG00701	C	Energy production and conversion	UgpQ	Glycerophosphoryl diester phosphodiesterase	COG00584	pfam03009	cd08568		1	2	1	1	1	1			2	1	1
arCOG00340	C	Energy production and conversion	GlcD	FAD/FMN-containing dehydrogenase fused to Heterodisulfide reductase, sub	COG00277	pfam01565,pfam02913,pfam13183		TIGR00387,TIGR	1		1		1	2					
arCOG00571	C	Energy production and conversion	SdhA	Succinate dehydrogenase/fumarate reductase, flavoprotein subunit	COG01053	pfam00890,pfam02910		TIGR03378,TIGR	2	1	1	1					1	2	8
arCOG00519	C	Energy production and conversion	FldA	Flavodoxin	COG00716	pfam12682			3	2	3	1	1	4					1
arCOG02410	C	Energy production and conversion	-	Coenzyme F420-dependent N5,N10-methylene tetrahydromethanopterin red	COG02141	pfam00296	cd01097	TIGR03555		1				1					2
arCOG01359	C	Energy production and conversion	-	Radical SAM superfamily enzyme	COG01031	pfam04055		TIGR04013									1	1	1
arCOG00338	C	Energy production and conversion	SdhC	Succinate dehydrogenase subunit C	COG02048	pfam02754,pfam02754		TIGR03288									1	1	2
arCOG03363	C	Energy production and conversion	NtpF	Archaeal/vacuolar-type H+-ATPase subunit H	COG02811												1		1
arCOG00024	C	Energy production and conversion	GlpK	Glycerol kinase	COG00554	pfam00370,pfam0278	cd07786	TIGR01311									1		
arCOG00335	C	Energy production and conversion	LutB	L-lactate utilization protein LutB, contains ferredoxin domain	COG01139	pfam02589,pfam13183		TIGR00273									1		
arCOG00349	C	Energy production and conversion	Fer	Ferredoxin	COG01141	pfam13459											1		
arCOG00509	C	Energy production and conversion	NorV	Flavorubredoxin	COG00426	pfam00753,pfam00258											1		
arCOG00709	C	Energy production and conversion	-	Aldehyde:ferredoxin oxidoreductase	COG02414	pfam02730,pfam01314											1		
arCOG00853	C	Energy production and conversion	SfcA	Malic enzyme	COG00281	pfam00390,pfam0394	cd05311										1		
arCOG00959	C	Energy production and conversion	-	Ferredoxin	COG01146			TIGR02512									1		
arCOG00964	C	Energy production and conversion	HdrC	Heterodisulfide reductase, subunit C	COG01150	pfam13183		TIGR03290									1		
arCOG01097	C	Energy production and conversion	-	Rubrerythrin	COG01592	pfam02915	cd01041										1		
arCOG01338	C	Energy production and conversion	-	Acyl-CoA synthetase, ATP-grasp containing subunit	COG01042	pfam13549											1		
arCOG01356	C	Energy production and conversion	-	Radical SAM superfamily enzyme	COG01032	pfam02310,pfam0405	cd02068,cd013	TIGR02026									1		
arCOG01549	C	Energy production and conversion	FrhA	Coenzyme F420-reducing hydrogenase, alpha subunit	COG03259	pfam00374		TIGR03295									1		
arCOG01552	C	Energy production and conversion	HycE	Ni,Fe-hydrogenase III component G	COG03262	pfam00329		TIGR01961									1		
arCOG01607	C	Energy production and conversion	PorA	Pyruvate:ferredoxin oxidoreductase or related 2-oxoacid:ferredoxin oxidored	COG00674	pfam01855	cd07034	TIGR03710									1		
arCOG01609	C	Energy production and conversion	-	Indolepyruvate ferredoxin oxidoreductase, alpha and beta subunit	COG04231	pfam01855,pfam0277	cd07034,cd020	TIGR03336									1		
arCOG01674	C	Energy production and conversion	AcyP	Acylphosphatase	COG01254	pfam00708											1		
arCOG01711	C	Energy production and conversion	Ppa	Inorganic pyrophosphatase	COG00221	pfam00719	cd00412										1		
arCOG02146	C	Energy production and conversion	-	Desulfoferredoxin	COG02033	pfam01880	cd03172	TIGR00332									1		
arCOG02428	C	Energy production and conversion	CdhA	CO dehydrogenase/acetyl-CoA synthase alpha subunit	COG01152	pfam03063	cd01916	TIGR00314									1		
arCOG02461	C	Energy production and conversion	-	Ferredoxin	COG01146	pfam13237	cd07030	TIGR03224									1		
arCOG02472	C	Energy production and conversion	FrhG	Coenzyme F420-reducing hydrogenase, gamma subunit	COG01941	pfam01058		TIGR03294									1		
arCOG02473	C	Energy production and conversion	FrhG	Coenzyme F420-reducing hydrogenase, gamma subunit	COG01941	pfam01058,pfam12838		TIGR03294									1		
arCOG02475	C	Energy production and conversion	FrhD	Coenzyme F420-reducing hydrogenase, delta subunit	COG01908	pfam02662											1		
arCOG02653	C	Energy production and conversion	FrhB	Coenzyme F420-reducing hydrogenase, beta subunit (with C-terminal ferredo	COG01035	pfam04432,pfam13183											1		
arCOG03626	C	Energy production and conversion	HyfE	Hydrogenase-4 membrane subunit HyfE	COG04237												1		
arCOG04002	C	Energy production and conversion	-	Cytochrome b family protein	SC.00119												1		

arCOG04391	C	Energy production and conversion	-	Rubredoxin	COG01773	pfam00301	cd00730											1				
arCOG04406	C	Energy production and conversion	FumA	Tartrate dehydratase beta subunit/Fumarate hydratase class I, C-terminal dor	COG01838	pfam05683		TIGR00723										1				
arCOG04407	C	Energy production and conversion	TtdA	Tartrate dehydratase alpha subunit/Fumarate hydratase class I, N-terminal do	COG01951	pfam05681		TIGR00722										1				
arCOG04429	C	Energy production and conversion	HyaD	Ni,Fe-hydrogenase maturation factor	COG00680	pfam01750	cd06067	TIGR00142										1				
arCOG04475	C	Energy production and conversion	OadB	Na+-transporting methylmalonyl-CoA/oxaloacetate decarboxylase, beta subu	COG01883	pfam03977		TIGR01109										1				
arCOG04537	C	Energy production and conversion	NuoF	NADH:ubiquinone oxidoreductase, NADH-binding 51 kD subunit (chain F)	COG01894	pfam01512,pfam10589		TIGR01959										1				
arCOG04874	C	Energy production and conversion	AllD	Malate/lactate/ureidoglycolate dehydrogenase, LDH2 family	COG02055	pfam02615		TIGR03175										1				
arCOG04890	C	Energy production and conversion	NuoE	NADH:ubiquinone oxidoreductase 24 kD subunit	COG01905	pfam01257	cd13637,cd030	TIGR01958										1				
arCOG05128	C	Energy production and conversion	NapF	Ferredoxin	COG01145	pfam13746		TIGR02910										1				
arCOG05744	C	Energy production and conversion	HycB	Fe-S-cluster-containing hydrogenase component 2	COG01142	pfam12838		TIGR01971										1				
arCOG05745	C	Energy production and conversion	-	BFD-like [2Fe-2S] binding domain	COG00446	pfam04324		TIGR01372										1				
arCOG05797	C	Energy production and conversion	OadG	Na+-transporting methylmalonyl-CoA/oxaloacetate decarboxylase, gamma su	COG03630	pfam04277		TIGR01195										1				
arCOG05865	C	Energy production and conversion	PepCK	Phosphoenolpyruvate carboxykinase, GTP-dependent	COG01274	pfam00821	cd00819											1				
arCOG06124	C	Energy production and conversion	ACH1	Acetyl-CoA hydrolase	COG00427	pfam02550,pfam13336		TIGR03458										1				
arCOG06130	C	Energy production and conversion	PflD	Pyruvate-formate lyase	COG01882	pfam02901,pfam0122	cd01677	TIGR01774										1				
arCOG13546	C	Energy production and conversion	-	ACP (acyl carrier) superfamily protein		pfam06857		TIGR01608										1				
arCOG01605	C	Energy production and conversion	PorD	Pyruvate:ferredoxin oxidoreductase or related 2-oxoacid:ferredoxin oxidored	COG01144	pfam12838		TIGR02179										2	1	1		
arCOG01163	C	Energy production and conversion	LeuB	Isocitrate/isopropylmalate dehydrogenase	COG00473	pfam00180		TIGR02088										2	1	3		
arCOG01340	C	Energy production and conversion	-	Acyl-CoA synthetase (NDP forming)	COG01042	pfam13380,pfam13607		TIGR02717										2				
arCOG01357	C	Energy production and conversion	-	Radical SAM superfamily enzyme	COG01032	pfam04055		TIGR04013										2				
arCOG01547	C	Energy production and conversion	HycE	Ni,Fe-hydrogenase III large subunit and subunit G	COG03261	pfam00374,pfam00346		TIGR01962										2				
arCOG01553	C	Energy production and conversion	-	Ni,Fe-hydrogenase III small subunit	COG03260	pfam01058		TIGR01957										2				
arCOG01601	C	Energy production and conversion	PorB	Pyruvate:ferredoxin oxidoreductase or related 2-oxoacid:ferredoxin oxidored	COG01013	pfam02775	cd03376	TIGR02176										2				
arCOG01602	C	Energy production and conversion	PorG	Pyruvate:ferredoxin oxidoreductase or related 2-oxoacid:ferredoxin oxidored	COG01014	pfam01558		TIGR03334										2				
arCOG01603	C	Energy production and conversion	PorG	Pyruvate:ferredoxin oxidoreductase or related 2-oxoacid:ferredoxin oxidored	COG01014	pfam01558		TIGR02175										2				
arCOG01608	C	Energy production and conversion	PorA	Pyruvate:ferredoxin oxidoreductase or related 2-oxoacid:ferredoxin oxidored	COG00674	pfam01855,pfam0278	cd07034	TIGR02176										2				
arCOG02235	C	Energy production and conversion	HdrA	Heterodisulfide reductase, subunit A or related polyferredoxin	COG01148	pfam07992,pfam13183,pfam01266,pf		TIGR03997,TIGR02179,TIGR03143,TIGR02179										2				
arCOG01545	C	Energy production and conversion	HyfC	Formate hydrogenlyase subunit 4	COG00650	pfam00146												3				
arCOG00706	C	Energy production and conversion	-	Aldehyde:ferredoxin oxidoreductase	COG02414	pfam02730,pfam01314												4				
arCOG01537	C	Energy production and conversion	NuoM	NADH:ubiquinone oxidoreductase subunit 4 (chain M)	COG00651	pfam00361		TIGR01972										5				
arCOG00855	C	Energy production and conversion	Pta	Phosphotransacetylase	COG00280	pfam01515		TIGR02706											1	1		
arCOG01167	C	Energy production and conversion	CoxL	Aerobic-type carbon monoxide dehydrogenase, large subunit CoxL/CutL hom	COG01529	pfam01315,pfam02738		TIGR02416											1	1		
arCOG01925	C	Energy production and conversion	CoxS	Aerobic-type carbon monoxide dehydrogenase, small subunit CoxS/CutS hom	COG02080	pfam13085,pfam0179	cd00207	TIGR03193											1	1		
arCOG01926	C	Energy production and conversion	CoxM	Aerobic-type carbon monoxide dehydrogenase, middle subunit CoxM/CutM h	COG01319	pfam00941,pfam03450		TIGR03195											1	1		
arCOG00962	C	Energy production and conversion	FrdB	Succinate dehydrogenase/fumarate reductase, Fe-S protein subunit	COG00479	pfam13085,pfam13183		TIGR00384											1	2		
arCOG04595	C	Energy production and conversion	QcrA	Rieske Fe-S protein	COG00723			TIGR01416											1	2		
arCOG00456	C	Energy production and conversion	GpsA	NAD(P)H-dependent glycerol-3-phosphate dehydrogenase	COG00240	pfam01210,pfam07479		TIGR03376											1	3		
arCOG01617	C	Energy production and conversion	Tas	Aryl-alcohol dehydrogenase related enzyme	COG00667	pfam00248	cd06660	TIGR01293											1	3		
arCOG04358	C	Energy production and conversion	FdhD	Uncharacterized protein required for formate dehydrogenase activity	COG01526	pfam02634		TIGR00129											1	3		
arCOG01236	C	Energy production and conversion	CyoA	Heme/copper-type cytochrome/quinol oxidase, subunit 2	COG01622	pfam00116	cd13842	TIGR02866											1			
arCOG01502	C	Energy production and conversion	HycB	Fe-S-cluster-containing hydrogenase component 2	COG01142	pfam13247		TIGR02512,TIGR03149											1			
arCOG02304	C	Energy production and conversion	Mct/CaiB	Succinyl-CoA:mesaconate CoA-transferase or predicted acyl-CoA transferase/	COG01804	pfam02515		TIGR03253											1			
arCOG02476	C	Energy production and conversion	HdrA	Heterodisulfide reductase, subunit A; ferredoxin domain	COG01148	pfam07992,pfam13237,pfam13454,pf		TIGR03140,TIGR04105,TIGR03385,TIGR04105											1			
arCOG04522	C	Energy production and conversion	CcdA	Cytochrome c biogenesis protein	COG00785														1			
arCOG02189	C	Energy production and conversion	NapF	HTH containing ranscriptional regulator fused to ferredoxin domain	COG01145			TIGR02179												1		
arCOG02842	C	Energy production and conversion	Fdx	Ferredoxin	COG00633	pfam00111	cd00207	TIGR02008												1		
arCOG04145	P	Inorganic ion transport and metabolism	TrkG	Trk-type K+ transport system, membrane component	COG00168	pfam02386		TIGR00933	1	1	1	1		1	1	1	1	1	1	1		
arCOG01959	P	Inorganic ion transport and metabolism	TrkA	TrkA, K+ transport system, NAD-binding component	COG00569	pfam02254,pfam02080,pfam02254,pfam02080			1	1	1	1		1	1	1		1	1			
arCOG04147	P	Inorganic ion transport and metabolism	SodA	Superoxide dismutase	COG00605	pfam00081,pfam02777			1	1	1				1	1						
arCOG04231	P	Inorganic ion transport and metabolism	CutA	Uncharacterized protein involved in tolerance to divalent cations	COG01324	pfam03091			1	1					1	1			1	3		
arCOG02021	P	Inorganic ion transport and metabolism	PspE	Rhodanese-related sulfurtransferase	COG00607	pfam00581	cd00158	TIGR02981	1		1				1	2			5	6		
arCOG07775	P	Inorganic ion transport and metabolism	HcaE	Phenylpropionate dioxygenase or related ring-hydroxylating dioxygenase, larg	COG04638	pfam00355,pfam0084	cd03535,cd088	TIGR03229	1	1	1	1	1		1				1	4		
arCOG02763	P	Inorganic ion transport and metabolism	-	Heavy-metal-associated domain (HMA)	COG02608	pfam00403	cd00371	TIGR02052	2	1					1				1	2		
arCOG02881	P	Inorganic ion transport and metabolism	ECM27	Ca2+/Na+ antiporter	COG00530	pfam01699,pfam01699		TIGR00367		1	1	1			1		1	2	4			
arCOG01477	P	Inorganic ion transport and metabolism	CzcD	Co/Zn/Cd efflux system component	COG01230	pfam01545		TIGR01297		2		1	1		1							
arCOG00576	P	Inorganic ion transport and metabolism	-	Predicted divalent heavy-metal cations transporter	COG00428	pfam02535					1		1		2	1						
arCOG00238	P	Inorganic ion transport and metabolism	ArsB	Na+/H+ antiporter NhaD or related arsenite permease	COG01055	pfam03600	cd01117	TIGR00785	2	1	2	1	1		2	2	3	1	2			
arCOG02569	P	Inorganic ion transport and metabolism	EriC	Chloride channel protein EriC	COG00038	pfam00654,pfam0057	cd00400,cd045	TIGR01302	1	2	1	1	1	1	2							
arCOG04330	P	Inorganic ion transport and metabolism	FTR1	High-affinity Fe2+/Pb2+ permease	COG00672	pfam03239		TIGR00145							2							
arCOG09746	P	Inorganic ion transport and metabolism	-	Sulfotransferase related protein		pfam00685			1	2			1			1				3		
arCOG04750	P	Inorganic ion transport and metabolism	-	Sirohydrochlorin iron chelatase fused to [2Fe-2S] Ferredoxin	COG02138	pfam01903,pfam0190	cd03416,cd03414,cd02980		1		1		1	1		1			1	1		
arCOG02499	P	Inorganic ion transport and metabolism	-	Lipoprotein NosD family, contains CASH domains	COG03420	pfam13229,pfam13229		TIGR04247	1							1						
arCOG02265	P	Inorganic ion transport and metabolism	CorA	Mg2+ and Co2+ transporter	COG00598	pfam01544	cd12828	TIGR00383								1	1					
arCOG04397	P	Inorganic ion transport and metabolism	AmtB	Ammonia permease	COG00004	pfam00909		TIGR00836								1	1					
arCOG02267	P	Inorganic ion transport and metabolism	PitA	Phosphate/sulphate permease	COG00306	pfam01384										1	2	1	1			
arCOG02640	P	Inorganic ion transport and metabolism	-	Uncharacterized protein YkaA, distantly related to PhoU, UPF0111/DUF47 fan	COG01392											1	2					
arCOG02764	P	Inorganic ion transport and metabolism	CopZ	Copper-ion-binding protein	COG02608	pfam00403	cd00371	TIGR02052								1		1	1			
arCOG01474	P	Inorganic ion transport and metabolism	MMT1	Predicted Co/Zn/Cd cation transporter	COG00053	pfam01545		TIGR01297								1						
arCOG02068	P	Inorganic ion transport and metabolism	DsrE	Uncharacterized protein involved in intracellular sulfur reduction	COG01553	pfam02635			1	1		1										
arCOG00173	P	Inorganic ion transport and metabolism	PhnE	ABC-type phosphate/phosphonate transport system, permease component	COG03639	pfam00528	cd06261	TIGR01097	1		1			1					1	3		
arCOG00206	P	Inorganic ion transport and metabolism	PhnC	ABC-type phosphate/phosphonate transport system, ATPase component	COG03638	pfam00005	cd03256	TIGR02315	1		1			1					1	3		
arCOG01805	P	Inorganic ion transport and metabolism	PhnD	ABC-type phosphate/phosphonate transport system, periplasmic component	COG03221	pfam12974	cd01071	TIGR01098	1		1			1					1	4		
arCOG04233	P	Inorganic ion transport and metabolism	FepB	ABC-type Fe3+-hydroxamate transport system, periplasmic component	COG00614	pfam01497	cd01143	TIGR04281	1		1								1			
arCOG00201	P	Inorganic ion transport and metabolism	ZnuC	ABC-type Mn/Zn transport system, ATPase component	COG01121	pfam00005	cd03235	TIGR03771	1			1						1	1	1		
arCOG01006	P	Inorganic ion transport and metabolism	ZnuB	ABC-type Mn2+/Zn2+ transport system, permease component	COG01108	pfam00950	cd06550	TIGR03770	1			1						1	1	1		
arCOG01005	P	Inorganic ion transport and metabolism	LraI	ABC-type metal ion transport system, periplasmic component/surface adhesi	COG00803	pfam01297	cd01018	TIGR03772,TIGR	1			1						1	1			
arCOG00163	P	Inorganic ion transport and metabolism	ThiP	ABC-type Fe3+ transport system, permease component	COG01178	pfam00528,pfam0052	cd06261,cd062	TIGR03262	1			1							1			
arCOG03127	P	Inorganic ion transport and metabolism	Kch	Ion channel fused to ion transporting domain	COG01226				1									1				
arCOG02851	P	Inorganic ion transport and metabolism	{NirD}	Ferredoxin subunit of nitrite reductase or ring-hydroxylating dioxygenase	COG02146	pfam00355	cd03467		1													
arCOG08934	P	Inorganic ion transport and metabolism	-	Alkylhydroperoxidase family enzyme	COG02128	pfam02627			1													
arCOG04487	P	Inorganic ion transport and metabolism	KatG	Catalase (peroxidase I)	COG00376	pfam00141,pfam0014	cd00649,cd006	TIGR00198		1												

arCOG02497	P	Inorganic ion transport and metabolism	-	Lipoprotein NosD family, contains CASH domains	COG03420	pfam12708		TIGR04247			1						1			
arCOG01578	P	Inorganic ion transport and metabolism	MgtA	Cation transport ATPase	COG00474	pfam00690,pfam00122,pfam00702,pf		TIGR01647									1	1	1	
arCOG02050	P	Inorganic ion transport and metabolism	-	Sulfite exporter, TauE/SafE family	COG00730	pfam01925											1	1	2	
arCOG00202	P	Inorganic ion transport and metabolism	EcfA2	Energy-coupling factor transporter ATP-binding protein EcfA2	COG01122	pfam00005	cd03225	TIGR01166									1	1		
arCOG01040	P	Inorganic ion transport and metabolism	CysC	Adenylylsulfate kinase or related kinase	COG00529	pfam01583	cd02027	TIGR00455									1		1	
arCOG01096	P	Inorganic ion transport and metabolism	CCC1	Predicted Fe2+/Mn2+ transporter, VIT1/CCC1 family	COG01814	pfam02915,pfam0198	cd01044,cd02431										1		1	
arCOG00219	P	Inorganic ion transport and metabolism	ModA	ABC-type molybdate transport system, periplasmic component	COG00725	pfam13531	cd13540	TIGR03730									1			
arCOG00359	P	Inorganic ion transport and metabolism	FeoB	Fe2+ transport system protein B	COG00370	pfam02421,pfam0767	cd01879	TIGR00437									1			
arCOG00624	P	Inorganic ion transport and metabolism	MgtE2	Permease, similar to cation transporter	COG01824	pfam01769											1			
arCOG01069	P	Inorganic ion transport and metabolism	-	CoA-dependent NAD(P)H Sulfur Oxidoreductase	COG00446	pfam07992,pfam02852		TIGR03385									1			
arCOG01095	P	Inorganic ion transport and metabolism	Ftn	Ferritin	COG01528	pfam00210	cd01055										1			
arCOG01475	P	Inorganic ion transport and metabolism	MMT1	Predicted Co/Zn/Cd cation transporter	COG00053	pfam01545,pfam0257	cd00851	TIGR01297									1			
arCOG01953	P	Inorganic ion transport and metabolism	KefB	Kef-type K+ transport system, membrane component	COG00475	pfam00999		TIGR00932									1			
arCOG01954	P	Inorganic ion transport and metabolism	KefB	Kef-type K+ transport system, membrane component	COG00475	pfam00999		TIGR00932									1			
arCOG01957	P	Inorganic ion transport and metabolism	TrkA	TrkA, K+ transport system, NAD-binding component	COG00569	pfam02254,pfam02080											1			
arCOG02102	P	Inorganic ion transport and metabolism	FeoA	Fe2+ transport system protein A	COG01918	pfam04023											1			
arCOG02248	P	Inorganic ion transport and metabolism	CbiM	ABC-type Co2+ transport system, permease component	COG00310	pfam01891		TIGR00123									1			
arCOG03077	P	Inorganic ion transport and metabolism	MnhB	Multisubunit Na+/H+ antiporter, MnhB subunit	COG02111												1			
arCOG03159	P	Inorganic ion transport and metabolism	CbiM	ABC-type Co2+ transport system, permease component	COG00310	pfam13190											1			
arCOG04191	P	Inorganic ion transport and metabolism	MET3	ATP sulfurylase (sulfate adenylyltransferase)	COG02046	pfam14306,pfam0174	cd00517	TIGR00339									1			
arCOG04355	P	Inorganic ion transport and metabolism	TehA	Tellurite resistance protein or related permease	COG01275	pfam03595	cd09299	TIGR00816									1			
arCOG05758	P	Inorganic ion transport and metabolism	-	TrkA-C domain contaning protein	COG00490	pfam02080											1			
arCOG05908	P	Inorganic ion transport and metabolism	-	Ferritin-like domain	COG01633	pfam04454											1			
arCOG10427	P	Inorganic ion transport and metabolism	-	Chromate resistance protein	COG04275	pfam09828											1			
arCOG10942	P	Inorganic ion transport and metabolism	-	Sulfotransferase family protein													1			
arCOG01103	P	Inorganic ion transport and metabolism	-	Ferritin-like domain	COG01633	pfam02915	cd01045										2			
arCOG01958	P	Inorganic ion transport and metabolism	Kch	Kef-type K+ transport system, predicted NAD-binding component	COG01226	pfam07885,pfam02254											2			
arCOG01960	P	Inorganic ion transport and metabolism	TrkA	K+ transport system, NAD-binding component fused to Ion channel	COG01226	pfam02254,pfam02080,pfam02254,pfam02080											2			
arCOG01961	P	Inorganic ion transport and metabolism	NhaP	NhaP-type Na+/H+ and K+/H+ antiporter	COG00025	pfam00999		TIGR00831									2			
arCOG02190	P	Inorganic ion transport and metabolism	ACR3	Arsenite efflux pump ACR3 or related permease	COG00798	pfam01758		TIGR00832									2			
arCOG03072	P	Inorganic ion transport and metabolism	MnhC	Multisubunit Na+/H+ antiporter, MnhC subunit	COG01006	pfam00420		TIGR00941									2			
arCOG03078	P	Inorganic ion transport and metabolism	-	Predicted subunit of the Multisubunit Na+/H+ antiporter	COG01563	pfam13244											2			
arCOG03079	P	Inorganic ion transport and metabolism	MnhB	Multisubunit Na+/H+ antiporter, MnhB subunit	COG02111	pfam04039											2			
arCOG03082	P	Inorganic ion transport and metabolism	MnhG	Multisubunit Na+/H+ antiporter, MnhG subunit	COG01320	pfam03334		TIGR01300									2			
arCOG03099	P	Inorganic ion transport and metabolism	MnhE	Multisubunit Na+/H+ antiporter, MnhE subunit	COG01863	pfam01899		TIGR00942									2			
arCOG03121	P	Inorganic ion transport and metabolism	MnhF	Multisubunit Na+/H+ antiporter, MnhF subunit	COG02212	pfam04066											2			
arCOG00318	P	Inorganic ion transport and metabolism	PhoU	Phosphate uptake regulator	COG00704	pfam04014,pfam01895,pfam01895		TIGR02135										1	1	
arCOG01576	P	Inorganic ion transport and metabolism	ZntA	Cation transport ATPase	COG02217	pfam00403,pfam0012	cd00371,cd014	TIGR00003,TIGR01511										1	1	
arCOG02062	P	Inorganic ion transport and metabolism	TusA	TusA-related sulfurtransferase	COG00425	pfam01206	cd00291											1	1	
arCOG02287	P	Inorganic ion transport and metabolism	-	Putative selenium binding protein, beta/alpha-propeller fold	COG00393	pfam01906												1	1	
arCOG05356	P	Inorganic ion transport and metabolism	Fes	Enterochelin esterase or related enzyme	COG02382	pfam00756												1	1	
arCOG11907	P	Inorganic ion transport and metabolism	-	Fe-S metabolizm associated domain		pfam02657		TIGR03391										1	1	
arCOG00230	P	Inorganic ion transport and metabolism	-	Periplasmic molybdate-binding protein/domain	COG01910	pfam00126,pfam12727		TIGR00637										1		
arCOG02849	P	Inorganic ion transport and metabolism	ArsA	Oxyanion-translocating ATPase	COG00003	pfam02374	cd02035	TIGR00345										1		
arCOG04559	P	Inorganic ion transport and metabolism	EmrE	Membrane transporter of cations and cationic drugs	COG02076	pfam00893												1		
arCOG01964	P	Inorganic ion transport and metabolism	Kch	Kef-type K+ transport system, predicted NAD-binding component	COG01226	pfam07885												2		
arCOG06837	P	Inorganic ion transport and metabolism	-	Lipoprotein NosD family, contains CASH domains	COG03420	pfam05048		TIGR04247,TIGR03024										2		
arCOG00167	P	Inorganic ion transport and metabolism	PstC	ABC-type phosphate transport system, permease component	COG00573		cd06261	TIGR02138											1	
arCOG00168	P	Inorganic ion transport and metabolism	PstA	ABC-type phosphate transport system, permease component	COG00581		cd06261	TIGR00974											1	
arCOG00213	P	Inorganic ion transport and metabolism	PstS	ABC-type phosphate transport system, periplasmic component	COG00226	pfam12849	cd13565	TIGR00975											1	
arCOG00231	P	Inorganic ion transport and metabolism	PstB	ABC-type phosphate transport system, ATPase component	COG01117	pfam00005	cd03260	TIGR00972											1	
arCOG00232	P	Inorganic ion transport and metabolism	PhoU	Phosphate uptake regulator	COG00704			TIGR02135											1	
arCOG02555	P	Inorganic ion transport and metabolism	-	Lipoprotein NosD family, contains CASH domains	COG03420														1	
arCOG02852	P	Inorganic ion transport and metabolism	{NirD}	Ferredoxin subunit of nitrite reductase or ring-hydroxylating dioxygenase	COG02146	pfam00355	cd03528	TIGR02377											1	
arCOG03797	P	Inorganic ion transport and metabolism	-	Ferritin-like domain	COG01633	pfam05763													1	
arCOG06728	P	Inorganic ion transport and metabolism	-	Sulfite exporter, TauE/SafE family	COG00730	pfam01925													2	
arCOG01709	I	Lipid transport and metabolism	CaiA	Acyl-CoA dehydrogenase	COG01960	pfam02771,pfam0277	cd01158	TIGR03207	1	1		1			1	1				
arCOG04213	I	Lipid transport and metabolism	INO1	Myo-inositol-1-phosphate synthase	COG01260	pfam07994		TIGR03450	1	2	1	1	1	1	1	1				
arCOG06112	I	Lipid transport and metabolism	Acs	Acyl-coenzyme A synthetase/AMP-(fatty) acid ligase	COG00365	pfam00501	cd05943	TIGR01217	1	2	1				1	1				
arCOG02245	I	Lipid transport and metabolism	-	Cytidylyltransferase family enzyme	COG01836	pfam01940		TIGR00297	1		1	1		1	1	1		1	2	
arCOG04106	I	Lipid transport and metabolism	CdsA	CDP-diglyceride synthetase	COG00575	pfam01864			1		1				1	1	1	1	1	
arCOG01532	I	Lipid transport and metabolism	UppS	Undecaprenyl pyrophosphate synthase	COG00020	pfam01255	cd00475	TIGR00055	1			1			1	1	1	1	2	
arCOG08932	I	Lipid transport and metabolism	LCB5	Diacylglycerol kinase family enzyme	COG01597	pfam00781		TIGR03702		1	1	1			1	1		1	2	
arCOG00249	I	Lipid transport and metabolism	FadB	3-hydroxyacyl-CoA dehydrogenase, some fused to Enoyl-CoA hydratase	COG01250	pfam02737,pfam0072	cd06558	TIGR02437,TIGR	3	2	3	2	1	1	1	2		2	4	
arCOG13950	I	Lipid transport and metabolism	-	Phosphatidylglycerophosphatase A		pfam04608	cd06971		1	1	1	1	1		1					
arCOG00860	I	Lipid transport and metabolism	-	Isopentenyl phosphate kinase, enzyme of modified mevalonate pathway	COG01608	pfam00696	cd04241	TIGR02075	1	1	1	1			1		2	1	2	
arCOG01028	I	Lipid transport and metabolism	ERG12	Mevalonate kinase	COG01577	pfam00288		TIGR00549	1	1	1			1	1		1	1	3	
arCOG01843	I	Lipid transport and metabolism	-	Sterol carrier protein	COG03255	pfam02036			1	1	1				1		1			
arCOG01264	I	Lipid transport and metabolism	FabG	Short-chain alcohol dehydrogenase	COG01028	pfam00106	cd05327	TIGR01289	1	1		1			1			1	1	
arCOG01085	I	Lipid transport and metabolism	PcrB	(S)-3-O-geranylgeranylglyceryl phosphate synthase, TIM-barrel fold	COG01646	pfam01884	cd02812	TIGR01768	1	1				1	1		1	1	4	
arCOG04351	I	Lipid transport and metabolism	-	Predicted membrane associated lipid hydrolase, neutral ceramidase superfam	COG03356	pfam09843			1	1				1	1		1			
arCOG00250	I	Lipid transport and metabolism	FadB	3-hydroxyacyl-CoA dehydrogenase	COG01250	pfam02737,pfam00725		TIGR02279	1		1	1	1	1	1			2	5	
arCOG04260	I	Lipid transport and metabolism	HMG1	Hydroxymethylglutaryl-CoA reductase	COG01257	pfam00368	cd00643	TIGR00533	1		1				1		1	1	2	
arCOG02936	I	Lipid transport and metabolism	ERG9	Phytoene/squalene synthetase	COG01562	pfam00494	cd00683	TIGR03465	2	2	1				1			3	5	
arCOG01710	I	Lipid transport and metabolism	Sbm	Methylmalonyl-CoA mutase, C-terminal domain/subunit (cobalamin-binding)	COG02185	pfam02310	cd02071	TIGR00640							1		1	1	3	
arCOG01590	I	Lipid transport and metabolism	AccC	Biotin carboxylase	COG00439	pfam00289,pfam02786,pfam02785		TIGR00514	1	1	1	1	1	1	2	1		2	7	
arCOG02705	I	Lipid transport and metabolism	-	Acetyl-CoA carboxylase, carboxyltransferase component	COG04799	pfam01039		TIGR01117	2	2		1	1	1	2	1	1	4	7	
arCOG04232	I	Lipid transport and metabolism	Sbm	Methylmalonyl-CoA mutase	COG01884	pfam01642	cd03680	TIGR00641	2	2		2			2	1	1	2	2	
arCOG01259	I	Lipid transport and metabolism	FabG	Short-chain alcohol dehydrogenase	COG01028		cd05233	TIGR01830	6	6	4	1	1	1	2	2	5	6	7	
arCOG01081	I	Lipid transport and metabolism	Idi	Isopentenyldiphosphate isomerase	COG01443	pfam00293	cd02885	TIGR02150	1	1	1	1		1	2					
arCOG00773	I	Lipid transport and metabolism	-	Acyl-CoA hydrolase	COG01607	pfam03061,pfam0306	cd03442,cd03442		1	1		1			2			1	3	

arCOG00670	I	Lipid transport and metabolism	PgsA	Phosphatidylglycerophosphate synthase	COG00558	pfam01066			1	2	2	2	1		2		1	2	2
arCOG00239	I	Lipid transport and metabolism	CaiD	Enoyl-CoA hydratase/carnithine racemase	COG01024	pfam00378	cd06558	TIGR03210	3		1	2		1	3	1		3	12
arCOG01707	I	Lipid transport and metabolism	CaiA	Acyl-CoA dehydrogenase	COG01960	pfam02771,pfam0277	cd00567	TIGR03207	4	3	3	2	1	1	3	1		4	11
arCOG04199	I	Lipid transport and metabolism	FAA1	Long-chain acyl-CoA synthetase (AMP-forming)	COG01022	pfam00501	cd05907	TIGR01923	1	1	1	2	1			1		1	2
arCOG01137	I	Lipid transport and metabolism	FadM	Acyl-CoA thioesterase FadM	COG00824	pfam13279	cd00586	TIGR00051	1		1	1				1		1	1
arCOG04100	I	Lipid transport and metabolism	CaiA	Acyl-CoA dehydrogenase	COG01960											1			
arCOG01529	I	Lipid transport and metabolism	Acs	Acyl-coenzyme A synthetase/AMP-(fatty) acid ligase	COG00365	pfam00501	cd05966	TIGR02188	1	1				1		2			
arCOG01282	I	Lipid transport and metabolism	PaaJ	Acetyl-CoA acetyltransferase	COG00183	pfam00108,pfam0280	cd00751	TIGR01930	2	2	2	1				2		3	5
arCOG01767	I	Lipid transport and metabolism	PksG	3-hydroxy-3-methylglutaryl CoA synthase	COG03425	pfam08545,pfam0854	cd00827	TIGR00748	1	1	1	1	1				1	3	2
arCOG01278	I	Lipid transport and metabolism	PaaJ	Acetyl-CoA acetyltransferase	COG00183		cd00829	TIGR01930	1	1	1	1					1	1	
arCOG04761	I	Lipid transport and metabolism	UppP	Undecaprenyl pyrophosphate phosphatase	COG01968	pfam02673		TIGR00753	1	2	1	1	1	1					
arCOG01879	I	Lipid transport and metabolism	-	Dolichol kinase family protein	COG00170				1			1						1	
arCOG01880	I	Lipid transport and metabolism	SEC59	Dolichol kinase	COG00170				1								2		
arCOG00856	I	Lipid transport and metabolism	CaiC	Acyl-CoA synthetase (AMP-forming)/AMP-acid ligase II	COG00318	pfam00501	cd12119	TIGR01923	1										
arCOG09737	I	Lipid transport and metabolism	LicD	Phosphorylcholine metabolism protein LicD	COG03475	pfam04991			1										
arCOG04470	I	Lipid transport and metabolism	Psd	Phosphatidylserine decarboxylase	COG00688	pfam02666		TIGR00164	2	1	1	1	1					1	
arCOG00671	I	Lipid transport and metabolism	PssA	Phosphatidylserine synthase	COG01183	pfam01066		TIGR04217	2	1		1	1						
arCOG00242	I	Lipid transport and metabolism	CaiD	Enoyl-CoA hydratase/carnithine racemase	COG01024	pfam00378	cd06558	TIGR02280		1									
arCOG01650	I	Lipid transport and metabolism	PldB	Lysophospholipase, alpha-beta hydrolase superfamily	COG02267	pfam12697		TIGR03695									1		1
arCOG03056	I	Lipid transport and metabolism	PgpB	Membrane-associated phospholipid phosphatase	COG00671	pfam01569	cd03392										1		
arCOG07155	I	Lipid transport and metabolism	IspD	4-diphosphocytidyl-2-methyl-D-erithritol synthase	COG01211	pfam01128	cd02516	TIGR00453									1		
arCOG11863	I	Lipid transport and metabolism	-	Membrane diacylglycerol kinase	COG00818	pfam01219	cd14263										1		
arCOG02228	I	Lipid transport and metabolism	GtrA	GtrA-like flippase	COG02246	pfam04138											2		
arCOG02039	I	Lipid transport and metabolism	Cls	Phosphatidylserine/phosphatidylglycerophosphate/cardiolipin synthase or re	COG01502	pfam13091,pfam1309	cd07493,cd09127,cd09128										3	1	
arCOG00241	I	Lipid transport and metabolism	CaiD	Enoyl-CoA hydratase/carnithine racemase	COG01024	pfam00378	cd06558	TIGR02280										1	1
arCOG01261	I	Lipid transport and metabolism	FabG	Short-chain alcohol dehydrogenase	COG01028	pfam00106	cd05233	TIGR01830										2	1
arCOG03951	I	Lipid transport and metabolism	PgpB	Membrane-associated phospholipid phosphatase	COG00671	pfam01569	cd03386											2	6
arCOG00674	I	Lipid transport and metabolism	PgsA	Phosphatidylglycerophosphate synthase	COG00558	pfam01066													1
arCOG06122	I	Lipid transport and metabolism	Acs	Acyl-coenzyme A synthetase/AMP-(fatty) acid ligase	COG00365	pfam00501	cd05959	TIGR02262											1
arCOG01487	F	Nucleotide transport and metabolism	ComEB	Deoxycytidylate deaminase	COG02131	pfam00383	cd01286	TIGR02571	1	1	1	1			1	1	1	1	1
arCOG00067	F	Nucleotide transport and metabolism	PrsA	Phosphoribosylpyrophosphate synthetase	COG00462	pfam13793,pfam0015	cd06223	TIGR01251	1	1	1				1	1	1	1	1
arCOG03214	F	Nucleotide transport and metabolism	ThyA	Thymidylate synthase	COG00207	pfam00303	cd00351	TIGR03283	1	1		1	1		1	1		1	4
arCOG02824	F	Nucleotide transport and metabolism	PurH	AICAR transformylase/IMP cyclohydrolase PurH	COG00138	pfam02142,pfam0180	cd01421	TIGR00355	1	2	1	1	1	1	1	1	1	1	2
arCOG00689	F	Nucleotide transport and metabolism	AllB	Dihydroorotase or related cyclic amidohydrolase	COG00044	pfam13147	cd01318	TIGR00857	1	2	1	1	1		1	1	1	1	2
arCOG04276	F	Nucleotide transport and metabolism	NrdA	Ribonucleotide reductase, alpha subunit	COG00209	pfam03477,pfam0031	cd02888	TIGR02504	1	2	2	1	1	1	1	1	2	1	3
arCOG01039	F	Nucleotide transport and metabolism	AdkA	Archaeal adenylate kinase	COG02019				1	2				1	1	1	1		
arCOG04184	F	Nucleotide transport and metabolism	RdgB	Inosine/xanthosine triphosphate pyrophosphatase, all-alpha NTP-PPase famil	COG00127	pfam01725	cd00515	TIGR00042	1		1			1	1	1	1	1	4
arCOG00028	F	Nucleotide transport and metabolism	-	Orotate phosphoribosyltransferase homolog	COG00856	pfam00156	cd06223	TIGR02985,TIGR	1		1				1	1	1	1	1
arCOG00639	F	Nucleotide transport and metabolism	PurM	Phosphoribosylaminoimidazole (AIR) synthetase	COG00150	pfam00586,pfam0276	cd02196	TIGR00878	2	2	1	1	1	1	1	1	1	1	1
arCOG00029	F	Nucleotide transport and metabolism	PyrE	Orotate phosphoribosyltransferase	COG00461	pfam00156	cd06223	TIGR00336		1	1	1			1	1	1	1	1
arCOG00692	F	Nucleotide transport and metabolism	SsnA	Cytosine deaminase or related metal-dependent hydrolase	COG00402	pfam01979	cd01305	TIGR02967		1	1				1	1	2		
arCOG00419	F	Nucleotide transport and metabolism	Hit	HIT family hydrolase	COG00537	pfam01230	cd01275			1				1	1	1	1	1	4
arCOG01747	F	Nucleotide transport and metabolism	PurB	Adenylosuccinate lyase	COG00015	pfam00206,pfam1039	cd01360	TIGR00928			1				1	1	1	1	2
arCOG00102	F	Nucleotide transport and metabolism	PurL	Phosphoribosylformylglycinamidine (FGAM) synthase, glutamine amidotransf	COG00047	pfam13507	cd01740	TIGR01737	1	1				1	1	2	1	1	2
arCOG04313	F	Nucleotide transport and metabolism	Ndk	Nucleoside diphosphate kinase	COG00105	pfam00334	cd04413		1	1				1	1	2	1	1	
arCOG00063	F	Nucleotide transport and metabolism	PyrG	CTP synthase (UTP-ammonia lyase)	COG00504	pfam06418,pfam0011	cd03113,cd017	TIGR00337	1	1	1	1	1	1	1		1	1	3
arCOG04462	F	Nucleotide transport and metabolism	PurS	Phosphoribosylformylglycinamidine (FGAM) synthase, PurS component	COG01828	pfam02700		TIGR00302	1	1	1	1	1		1		1	1	2
arCOG04421	F	Nucleotide transport and metabolism	PurC	Phosphoribosylaminoimidazolesuccinocarboxamide (SAICAR) synthase	COG00152	pfam01259	cd01415	TIGR00081	1	1	1	1	1		1		1	1	
arCOG00641	F	Nucleotide transport and metabolism	PurL	Phosphoribosylformylglycinamidine (FGAM) synthase, synthetase domain	COG00046	pfam00586,pfam0276	cd02203,cd022	TIGR01736	1	1	1	1	1		1		1		2
arCOG01565	F	Nucleotide transport and metabolism	NrnA	nanoRNase/pAp phosphatase, hydrolyzes c-di-AMP and oligoRNAs	COG00618	pfam01368			1	1	1		1		1		1	1	1
arCOG00911	F	Nucleotide transport and metabolism	PyrB	Aspartate carbamoyltransferase, catalytic chain	COG00540	pfam02729,pfam00185		TIGR00670	1	2	1	1		1	1		1	1	2
arCOG01034	F	Nucleotide transport and metabolism	THEP1	Nucleoside-triphosphatase THEP1	COG01618	pfam03266	cd00009		1	2	1	1		1	1		1	1	2
arCOG04229	F	Nucleotide transport and metabolism	PyrI	Aspartate carbamoyltransferase, regulatory subunit	COG01781	pfam01948,pfam02748		TIGR00240	1	2	1	1		1	1		1	1	2
arCOG04048	F	Nucleotide transport and metabolism	Dcd	Deoxycytidine deaminase	COG00717	pfam00692	cd07557	TIGR02274	1		1				1		1	1	2
arCOG04387	F	Nucleotide transport and metabolism	PurA	Adenylosuccinate synthase	COG00104	pfam00709	cd03108	TIGR00184	1		1				1		1	1	2
arCOG04415	F	Nucleotide transport and metabolism	PurD	Phosphoribosylamine-glycine ligase	COG00151	pfam02844,pfam0107	1,pfam02843	TIGR00877		1	1	1			1		1	1	1
arCOG01038	F	Nucleotide transport and metabolism	Fap7	Broad-specificity NMP kinase	COG01936	pfam13238				1	1	1			1		1	1	2
arCOG00018	F	Nucleotide transport and metabolism	Nnr2	NAD(P)H-hydrate repair enzyme Nnr, NAD(P)H-hydrate dehydratase domain	COG00063	pfam03853,pfam0125	cd01171	TIGR00197,TIGR00196		1	1	1			1		1		
arCOG00085	F	Nucleotide transport and metabolism	GuaA	GMP synthase, PP-ATPase domain/subunit	COG00519	pfam02540,pfam0095	cd01997	TIGR00884		1	1			1	1		1	1	1
arCOG00090	F	Nucleotide transport and metabolism	GuaA	GMP synthase - Glutamine amidotransferase domain	COG00518	pfam00117	cd01741	TIGR00888		1					1				3
arCOG04346	F	Nucleotide transport and metabolism	-	5-formaminoimidazole-4-carboxamide-1-beta-D-ribofuranosyl 5'-monophosp	COG01759	pfam06849,pfam0697	3	TIGR00877							1			2	
arCOG00603	F	Nucleotide transport and metabolism	PyrD	Dihydroorotate dehydrogenase	COG00167	pfam01180	cd04740	TIGR01037	1	1	1	1	1	1	2	1	1	1	4
arCOG01037	F	Nucleotide transport and metabolism	Cmk	Cytidylate kinase	COG01102	pfam13189	cd02020	TIGR02173	1	2	1				2	1	1	1	1
arCOG00087	F	Nucleotide transport and metabolism	GuaA	GMP synthase - Glutamine amidotransferase domain	COG00518	pfam00117	cd01742	TIGR00888	2	1	1	2			2	1	1	1	1
arCOG01891	F	Nucleotide transport and metabolism	Tmk	Thymidylate kinase	COG00125	pfam02223	cd01672	TIGR00041	2	1		1		1	2	1	1	1	2
arCOG00093	F	Nucleotide transport and metabolism	PurF	Glutamine phosphoribosylpyrophosphate amidotransferase	COG00034	pfam00310,pfam0015	cd00715,cd062	TIGR01134	3	3				1	2	2	1	1	2
arCOG02825	F	Nucleotide transport and metabolism	PurN	Folate-dependent phosphoribosylglycinamide formyltransferase PurN	COG00299				3	2	1	1		2	3	2	1	2	3
arCOG02464	F	Nucleotide transport and metabolism	PurE	Phosphoribosylcarboxyaminoimidazole (NCAIR) mutase	COG00041	pfam00731		TIGR01162	1	1		1				1	1	1	1
arCOG00612	F	Nucleotide transport and metabolism	GuaB	IMP dehydrogenase/GMP reductase	COG00516	pfam00478	cd00381	TIGR01302	2	2		2				1	1		3
arCOG02807	F	Nucleotide transport and metabolism	AzgA	Xanthine/uracil/vitamin C permease, AzgA family	COG02252	pfam00860				1	1	1				1	1		2
arCOG05252	F	Nucleotide transport and metabolism	-	Predicted secreted endonuclease distantly related to archaeal Holliday junctio	COG04741	pfam10107										1			
arCOG00030	F	Nucleotide transport and metabolism	Apt	Adenine/guanine phosphoribosyltransferase or related PRPP-binding protein	COG00503	pfam00156	cd06223	TIGR01090	1	1		1				4	1	1	4
arCOG00081	F	Nucleotide transport and metabolism	PyrF	Orotidine-5'-phosphate decarboxylase	COG00284	pfam00215	cd04725	TIGR01740	1	2							1	1	2
arCOG03575	F	Nucleotide transport and metabolism	-	Polyphosphate kinase 2	COG02326	pfam03976,pfam0397	cd01672	TIGR03708	1			1		1					
arCOG00858	F	Nucleotide transport and metabolism	PyrH	Uridylate kinase	COG00528	pfam00696	cd04253	TIGR02076									1	1	1
arCOG04541	F	Nucleotide transport and metabolism	MIS1	Formyltetrahydrofolate synthetase	COG02759	pfam01268	cd00477										1	1	2
arCOG01350	F	Nucleotide transport and metabolism	-	Predicted inorganic polyphosphate/ATP-NAD kinase	COG03199	pfam01513											1	1	3
arCOG02013	F	Nucleotide transport and metabolism	DeoA	Thymidine phosphorylase	COG00213	pfam01568,pfam0288	5,pfam00591,pf	TIGR03327									1	1	
arCOG04311	F	Nucleotide transport and metabolism	-	5'-deoxynucleotidase, HD superfamily hydrolase	COG01896	pfam13023											1	1	
arCOG01075	F	Nucleotide transport and metabolism	-	NUDIX family hydrolase	COG01051	pfam00293	cd04673										1	4	8
arCOG04320	F	Nucleotide transport and metabolism	DeoC	Deoxyribose-phosphate aldolase	COG00274	pfam01791	cd00959	TIGR00126									1		1

arCOG01324	F	Nucleotide transport and metabolism	Udp	Uridine phosphorylase	COG02820	pfam01048		TIGR01718									1			
arCOG01723	F	Nucleotide transport and metabolism	CyaB	Adenylate cyclase, class 2 (thermophilic)	COG01437	pfam01928	cd07890	TIGR00318									1			
arCOG01883	F	Nucleotide transport and metabolism	THY1	Thymidylate synthase	COG01351	pfam02511		TIGR02170									1			
arCOG04173	F	Nucleotide transport and metabolism	Cdd	Cytidine deaminase	COG00295	pfam00383	cd01283	TIGR01354									1			
arCOG04298	F	Nucleotide transport and metabolism	-	Adenosine/AMP kinase	COG01839	pfam04008											1			
arCOG04309	F	Nucleotide transport and metabolism	-	S-adenosyl-l-methionine hydroxide adenosyltransferase	COG01912	pfam01887											1			
arCOG04889	F	Nucleotide transport and metabolism	NrdD	Oxygen-sensitive ribonucleoside-triphosphate reductase	COG01328	pfam13597	cd01675	TIGR02487									1			
arCOG05133	F	Nucleotide transport and metabolism	Udk	Uridine kinase	COG00572	pfam00485	cd02026	TIGR00235									1			
arCOG01327	F	Nucleotide transport and metabolism	Pnp	Purine nucleoside phosphorylase	COG00005	pfam01048		TIGR01694									2	1	2	
arCOG00695	F	Nucleotide transport and metabolism	SsnA	Cytosine deaminase or related metal-dependent hydrolase	COG00402	pfam01979	cd01298	TIGR03314									2	2	2	
arCOG03658	F	Nucleotide transport and metabolism	NrdF	Ribonucleotide reductase, beta subunit (ferritin domain)	COG00208	pfam00268	cd01049	TIGR04171										1	1	
arCOG01046	F	Nucleotide transport and metabolism	Adk	Adenylate kinase or related kinase	COG00563	pfam00406	cd01428	TIGR01351										1	2	
arCOG14568	F	Nucleotide transport and metabolism	-	Pseudouridine synthase		pfam00849	cd02554	TIGR00093										1	2	
arCOG01566	F	Nucleotide transport and metabolism	NrnA	nanoRNase/pAp phosphatase, hydrolyzes c-di-AMP and oligoRNAs	COG00618	pfam02254,pfam01368,pfam02272												1		
arCOG06976	Q	Secondary metabolites biosynthesis, transport and	HmgA	Homogentisate 1,2-dioxygenase	COG03508				1	1			1	1	1	1				
arCOG00696	Q	Secondary metabolites biosynthesis, transport and	HutI	Imidazolonepropionase or related amidohydrolase	COG01228	pfam13147	cd01296	TIGR01224	1			1		1	1	1	2	1	2	
arCOG00235	Q	Secondary metabolites biosynthesis, transport and	MhpD	2-keto-4-pentenoate hydratase/2-oxohepta-3-ene-1,7-dioic acid hydratase (c	COG00179	pfam01557		TIGR02303	2	1		1	1	1	1	1	1	1	1	
arCOG04347	Q	Secondary metabolites biosynthesis, transport and	-	SAM-dependent methyltransferase	COG00500	pfam08241	cd02440	TIGR01934	3	2	1	1		2	1					
arCOG01730	Q	Secondary metabolites biosynthesis, transport and	-	Aromatic ring-opening dioxygenase, catalytic LigB subunit related enzyme	COG03384	pfam02900	cd07363	TIGR02298							1					
arCOG04340	Q	Secondary metabolites biosynthesis, transport and	-	SAM-dependent methyltransferase	COG00500	pfam13489	cd02440	TIGR02081							1					
arCOG06106	Q	Secondary metabolites biosynthesis, transport and	-	Predicted ring-cleavage extradiol dioxygenase	COG02514	pfam12681,pfam1268	cd07255,cd072	TIGR03211								1				
arCOG03570	Q	Secondary metabolites biosynthesis, transport and	-	SAM-dependent methyltransferase	COG00500	pfam12847	cd02440	TIGR02021	1	1	1	1		1					1	
arCOG01791	Q	Secondary metabolites biosynthesis, transport and	-	SAM-dependent methyltransferase	COG00500	pfam12847	cd02440	TIGR02021	1	1		1	1	1			1			
arCOG01792	Q	Secondary metabolites biosynthesis, transport and	-	SAM-dependent methyltransferase	COG00500	pfam08241	cd02440	TIGR01934	1	1		1								
arCOG01781	Q	Secondary metabolites biosynthesis, transport and	-	SAM-dependent methyltransferase	COG00500	pfam08241	cd02440	TIGR02072	1	1			1	1				1	2	
arCOG00777	Q	Secondary metabolites biosynthesis, transport and	PaaI	HGG motif-containing thioesterase, possibly involved in aromatic compounds	COG02050	pfam03061	cd03443	TIGR00369	1	2	1	1	1					1	2	
arCOG01521	Q	Secondary metabolites biosynthesis, transport and	-	Phytoene dehydrogenase or related enzyme	COG01233	pfam13450			1	2				1				1	3	
arCOG03914	Q	Secondary metabolites biosynthesis, transport and	SufI	Multicopper oxidase	COG02132	pfam07732	cd11024	TIGR02376	1		1		1							
arCOG01523	Q	Secondary metabolites biosynthesis, transport and	-	Phytoene dehydrogenase or related enzyme	COG01233	pfam01593		TIGR03467	1		1			1				1	1	
arCOG01778	Q	Secondary metabolites biosynthesis, transport and	-	SAM-dependent methyltransferase	COG00500	pfam08241	cd02440	TIGR02072		1	1		1	1						
arCOG01782	Q	Secondary metabolites biosynthesis, transport and	-	SAM-dependent methyltransferase	COG00500	pfam08241	cd02440	TIGR01934			1									
arCOG05015	Q	Secondary metabolites biosynthesis, transport and	-	SAM-dependent methyltransferase	COG00500	pfam12847	cd02440	TIGR03534									1		1	
arCOG01229	Q	Secondary metabolites biosynthesis, transport and	-	Cyclic 2, 3-diphosphoglycerate synthetase	COG02403												1			
arCOG01400	Q	Secondary metabolites biosynthesis, transport and	-	SAM-dependent methyltransferase	COG00500	pfam05050	cd02440,cd021	TIGR01444									1			
arCOG01729	Q	Secondary metabolites biosynthesis, transport and	-	Aromatic ring-opening dioxygenase, LigB subunit	COG03885	pfam02900	cd07952										1			
arCOG01773	Q	Secondary metabolites biosynthesis, transport and	-	SAM-dependent methyltransferase	COG00500	pfam08241		TIGR01934									1			
arCOG02702	Q	Secondary metabolites biosynthesis, transport and	-	SAM-dependent methyltransferase	COG00500	pfam08241	cd02440	TIGR02072									1			
arCOG03688	Q	Secondary metabolites biosynthesis, transport and	osmC	Organic hydroperoxide reductase	COG01764	pfam02566		TIGR03561									1			
arCOG01943	Q	Secondary metabolites biosynthesis, transport and	PncA	Amidase related to nicotinamidase	COG01335	pfam00857	cd00431	TIGR03614									2	1		
arCOG04786	Q	Secondary metabolites biosynthesis, transport and	-	1,2-phenylacetyl-CoA epoxidase, catalytic subunit	COG03396	pfam05138		TIGR02158										1	3	
arCOG06169	Q	Secondary metabolites biosynthesis, transport and	-	Arylsulfotransferase family protein		pfam05935												1		
arCOG01402	Q	Secondary metabolites biosynthesis, transport and	AglP	SAM-dependent methyltransferase	COG00500	pfam05050		TIGR01444										2	2	
POORLY CHARACTERIZED																				
arCOG02177	S	Function unknown	-	Uncharacterized membrane protein	COG01967	pfam01889			1	1	1	1	1	1	1	1		1	2	
arCOG04565	S	Function unknown	-	Predicted membrane protein, DUF368 family	COG02035	pfam04018			1	1	1	1		1	1	1				
arCOG05517	S	Function unknown	-	Uncharacterized protein					1	1	1	1			1	1	1		1	
arCOG04619	S	Function unknown	-	Uncharacterized membrane protein, contains bPH2 (bacterial pleckstrin homo	COG03428	pfam03703,pfam03703,pfam03703			1	1	1	1			1	1			1	
arCOG05368	S	Function unknown	-	Uncharacterized protein	SC.00138				1	1		1	1	1	1	1			1	
arCOG04521	S	Function unknown	-	Uncharacterized membrane protein, predicted permease	SC.00496				1		1	1		1	1	1	1	1	2	
arCOG04596	S	Function unknown	-	Uncharacterized membrane protein					1		1			1	1	1				
arCOG03678	S	Function unknown	-	Uncharacterized protein	SC.00067				1		1				1	1	1	1	3	
arCOG04308	S	Function unknown	-	Uncharacterized protein	COG01698	pfam03685			1			1			1	1	1	1	1	
arCOG03729	S	Function unknown	-	Metal-binding cluster containing protein						1					1	1				
arCOG13537	S	Function unknown	-	Uncharacterized protein					1	1	1		1	2	1		1			
arCOG11014	S	Function unknown	-	HTH-domain containing transcriptional regulator	SC.00234	pfam06224			1	1	1				1			1	1	
arCOG01224	S	Function unknown	-	Uncharacterized protein	COG01849	pfam04010			1	1		1			1		1	1	2	
arCOG02142	S	Function unknown	-	Pheromone shutdown protein TraB, contains GTxH motif	COG01916	pfam01963		TIGR00261	1	2	1	1			1		1	1	1	
arCOG03124	S	Function unknown	-	Pentapeptide repeats containing protein	COG01357	pfam00805,pfam00805			1	2	1	1			1			2	3	
arCOG03633	S	Function unknown	-	Uncharacterized membrane protein YckC, RDD family	COG01714	pfam06271			1	2					1		1			
arCOG10198	S	Function unknown	-	Uncharacterized membrane protein, DUF2899 family		pfam11449			1	2					1				1	
arCOG07412	S	Function unknown	-	Uncharacterized protein					1		1			1	1		1			
arCOG05495	S	Function unknown	-	Uncharacterized protein, contains N-terminal coiled-coil domain	COG04911	pfam09969			1			1			1					
arCOG04076	S	Function unknown	-	Uncharacterized protein, DUF359 family	COG01909	pfam04019			1						1		1	1	3	
arCOG07813	S	Function unknown	-	LamG-like jellyroll fold domain	SC.00184	pfam13385			1						1			1		
arCOG04370	S	Function unknown	-	Uncharacterized membrane protein	COG03503	pfam07786			1						1					
arCOG04412	S	Function unknown	-	Uncharacterized protein	COG04004					1	1	1			1		1	1		
arCOG08231	S	Function unknown	-	Periplasmic protein with immunoglobin-like fold		pfam00932				1					1		1	1		
arCOG05330	S	Function unknown	-	Uncharacterized protein						1					1		1			
arCOG02546	S	Function unknown	-	Secreted protein, contains PKD repeats and vWA domain	COG03291	pfam05048,pfam0504	cd00146,cd001	TIGR04247,TIGR04247,TI		1					1				2	
arCOG03606	S	Function unknown	-	Uncharacterized protein	SC.00401							1			1					
arCOG01917	S	Function unknown	-	Zn-ribbon domain containing protein											1		3		1	
arCOG02998	S	Function unknown	-	Cupin domain containing protein	COG01917	pfam07883									1			1	2	
arCOG02492	S	Function unknown	-	Secreted protein, with PKD repeat domain	COG03291										1					
arCOG02761	S	Function unknown	-	Uncharacterized protein, DUF302 family	COG03439	pfam03625	cd14797								1					
arCOG04321	S	Function unknown	-	Uncharacterized protein DUF711 family, similar to ribonucleotide reductase a	COG02848	pfam05167	cd08025								1					
arCOG05852	S	Function unknown	-	HEAT repeats containing protein	COG01413										1					
arCOG01159	S	Function unknown	-	Uncharacterized archaeal coiled-coil protein	COG01340			TIGR02168	2	2	2	1	1		2	1	2	2	5	
arCOG01314	S	Function unknown	-	Uncharacterized membrane anchored protein with extracellular flavodoxin-lik	SC.00272				2	1	1	1			2	2		1		
arCOG08673	S	Function unknown	-	Uncharacterized protein					1	1	1	1			2			1		

arCOG05351	S	Function unknown	-	Uncharacterized membrane protein		pfam11433				2					2		1		
arCOG01907	S	Function unknown	AIM24	Uncharacterized protein, AIM24 family	COG02013	pfam01987		TIGR00266	3	3	3	3		3	3	3	1		1
arCOG03232	S	Function unknown	-	VanZ like family protein	COG05652	pfam04892			1	1		1				1	1		
arCOG07442	S	Function unknown	-	Uncharacterized protein					1	1						1			
arCOG02087	S	Function unknown	-	Predicted membrane protein	COG01470	pfam10633,pfam13620,pfam10633			1	2		1		1		1			2
arCOG04214	S	Function unknown	-	Uncharacterized protein	COG00432	pfam01894		TIGR00149	1		1		1			1	1	1	
arCOG10338	S	Function unknown	-	Uncharacterized protein	SC.00914	pfam13517,pfam13517,pfam13517,pfam13517				2	1		1			1	1		
arCOG05338	S	Function unknown	-	Uncharacterized protein							1		1			1		1	2
arCOG02081	S	Function unknown	-	Predicted membrane protein	COG01470	pfam10633						1				1		2	2
arCOG02206	S	Function unknown	-	Uncharacterized membrane protein										1		1	1		1
arCOG01330	S	Function unknown	-	Uncharacterized membrane protein	COG02426	pfam06695										1	2	1	3
arCOG04269	S	Function unknown	-	Protein, predicted to be involved in DNA repair	COG01602	pfam04894,pfam04895										1	2		
arCOG02565	S	Function unknown	-	Secreted uncharacterized protein	COG05276											1			
arCOG06429	S	Function unknown	-	Uncharacterized protein, contains NRDE domain	COG03332	pfam05742										1			
arCOG02508	S	Function unknown	-	Secreted protein, with PKD repeat domain	COG03291	pfam00801,pfam0080	cd00146,cd001	TIGR00864	1	1	2			1		3		1	
arCOG02884	S	Function unknown	-	Membrane associated protein with extracellular Ig-like domain, a component	COG04743				1	1	1	1		1			2		
arCOG04618	S	Function unknown	-	Uncharacterized protein, DUF2237 family	COG03651	pfam09996			1	1		1	1					1	2
arCOG10444	S	Function unknown	-	Uncharacterized protein					1	1		1	1						4
arCOG12677	S	Function unknown	-	Uncharacterized protein					1	1			1	1					
arCOG04693	S	Function unknown	-	Uncharacterized protein					1	1									1
arCOG03888	S	Function unknown	-	Uncharacterized membrane protein					1	1									
arCOG03949	S	Function unknown	-	Uncharacterized membrane protein	COG03815	pfam09858			1	2	1	1	1	1			1		2
arCOG07680	S	Function unknown	-	Uncharacterized membrane protein	COG02832	pfam04304			1	2	1			1				1	3
arCOG10153	S	Function unknown	-	Uncharacterized membrane protein, DUF4112 family		pfam13430			1	2		1							
arCOG04364	S	Function unknown	-	Uncharacterized protein	COG01772	pfam04407			1	3	1	1	1	2					
arCOG02491	S	Function unknown	-	WD40 repeats containing protein	COG02319	pfam13360			1		1	1		2				1	
arCOG01119	S	Function unknown	-	GYD domain, alpha/beta barrel superfamily	COG04274	pfam08734			1		1	1							
arCOG08355	S	Function unknown	-	Uncharacterized membrane protein					1		1		1	1					
arCOG04566	S	Function unknown	-	Uncharacterized protein					1			1							
arCOG02966	S	Function unknown	-	HEAT repeats containing protein	COG01413				1				1	1					
arCOG03749	S	Function unknown	-	Uncharacterized membrane protein	COG04243	pfam07884	cd12918		1				1						
arCOG06533	S	Function unknown	-	Uncharacterized membrane protein					1					2					
arCOG05022	S	Function unknown	-	Uncharacterized protein, contains PQ loop repeat	COG04095				1									1	1
arCOG11882	S	Function unknown	-	Uncharacterized membrane protein					1										2
arCOG05839	S	Function unknown	-	Uncharacterized protein				TIGR04292	2	1	2		2	2					
arCOG08211	S	Function unknown	-	Uncharacterized membrane protein	SC.00448					1	1			1					
arCOG08977	S	Function unknown	-	Uncharacterized membrane protein	COG04270					1	1							1	2
arCOG09475	S	Function unknown	-	Uncharacterized protein	SC.00862					1									
arCOG06958	S	Function unknown	-	Uncharacterized protein						2							1		2
arCOG04006	S	Function unknown	-	HEAT repeats containing protein	COG01413	pfam13646				2								1	3
arCOG04579	S	Function unknown	-	Uncharacterized protein, DUF2071 family	COG03361	pfam09844					1		1	1				1	2
arCOG06742	S	Function unknown	-	NosD-like cell surface protein							1		1						
arCOG08643	S	Function unknown	-	Secreted protein with C-terminal PEFG domain							1								
arCOG02527	S	Function unknown	-	Secreted protein, with PKD repeat domain	COG03291							1					2		
arCOG04662	S	Function unknown	-	Uncharacterized membrane protein	SC.00128												1	1	
arCOG05345	S	Function unknown	-	Uncharacterized protein													1	1	
arCOG04500	S	Function unknown	-	Uncharacterized protein with Ig-like domain													1	2	1
arCOG02170	S	Function unknown	-	Uncharacterized protein	SC.00300	pfam13559											1		1
arCOG09426	S	Function unknown	-	Uncharacterized membrane protein													1		2
arCOG01302	S	Function unknown	-	Uncharacterized protein	COG01531	pfam04457											1		
arCOG01336	S	Function unknown	AMMECR1	Uncharacterized protein	COG02078	pfam01871		TIGR04335									1		
arCOG01472	S	Function unknown	-	Hemerythrin HHE cation binding domain containing protein	COG02461	pfam04282,pfam01814,pfam13596											1		
arCOG01921	S	Function unknown	-	Uncharacterized protein, DUF61 family	COG02083	pfam01886											1		
arCOG02148	S	Function unknown	-	Uncharacterized homolog of gamma-carboxymuconolactone decarboxylase su	COG00599	pfam02627		TIGR04169									1		
arCOG02159	S	Function unknown	-	Uncharacterized membrane protein	COG04089	pfam07758											1		
arCOG02264	S	Function unknown	-	Uncharacterized membrane protein	COG01808	pfam04087		TIGR00341									1		
arCOG02566	S	Function unknown	-	Uncharacterized membrane protein													1		
arCOG02717	S	Function unknown	-	Predicted membrane protein, DUF131 family	COG02034	pfam01998											1		
arCOG02991	S	Function unknown	-	Uncharacterized protein, DUF1850 family	COG04729	pfam08905											1		
arCOG03107	S	Function unknown	-	Uncharacterized protein, YigZ/IMPACT family	COG01739	pfam01205,pfam09186		TIGR00257									1		
arCOG03128	S	Function unknown	-	Pentapeptide repeats containing protein	COG01357	pfam13599,pfam13599											1		
arCOG03165	S	Function unknown	-	Uncharacterized membrane protein													1		
arCOG03379	S	Function unknown	-	Uncharacterized protein													1		
arCOG03397	S	Function unknown	-	Uncharacterized protein													1		
arCOG03573	S	Function unknown	-	Uncharacterized protein, DUF1015 family	COG04198	pfam06245											1		
arCOG03677	S	Function unknown	-	Uncharacterized protein	SC.00067												1		
arCOG03768	S	Function unknown	-	Uncharacterized membrane protein, DUF973 family	SC.00418	pfam06157											1		
arCOG03776	S	Function unknown	-	Zn finger protein	SC.00092												1		
arCOG03853	S	Function unknown	-	Uncharacterized membrane protein													1		
arCOG04051	S	Function unknown	-	Uncharacterized protein	COG02412	pfam04242											1		
arCOG04079	S	Function unknown	-	Uncharacterized protein													1		
arCOG04132	S	Function unknown	-	Uncharacterized protein	COG04697	pfam09910											1		
arCOG04140	S	Function unknown	-	Uncharacterized protein	COG01888	pfam02680											1		
arCOG04253	S	Function unknown	-	Uncharacterized protein	COG01415	pfam05559											1		
arCOG04361	S	Function unknown	-	Predicted metal-binding protein	COG05423	pfam10050											1		
arCOG04373	S	Function unknown	-	Uncharacterized protein YqgV, UPF0045/DUF77 family	COG00011	pfam01910		TIGR00106									1		
arCOG04390	S	Function unknown	-	Uncharacterized protein containing a Zn-ribbon	COG04068	pfam09889											1		
arCOG04424	S	Function unknown	-	Uncharacterized protein	COG03377	pfam08827											1		
arCOG04484	S	Function unknown	-	Uncharacterized membrane protein, DUF1648 family	COG05658	pfam07853,pfam13630											1		
arCOG04555	S	Function unknown	-	Uncharacterized membrane protein related to bactofilin		pfam04519,pfam04519											1		
arCOG04705	S	Function unknown	-	Uncharacterized protein	COG02098	pfam04036,pfam04038											1		

arCOG04938	S	Function unknown	-	Uncharacterized membrane protein	SC.00516													1		
arCOG05238	S	Function unknown	-	Uncharacterized protein														1		
arCOG05308	S	Function unknown	-	Uncharacterized membrane protein	SC.00137													1		
arCOG05325	S	Function unknown	-	Uncharacterized protein														1		
arCOG05327	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG05352	S	Function unknown	-	Uncharacterized protein														1		
arCOG05357	S	Function unknown	-	Uncharacterized protein														1		
arCOG05358	S	Function unknown	-	Uncharacterized protein														1		
arCOG05389	S	Function unknown	-	Uncharacterized protein														1		
arCOG05460	S	Function unknown	-	Uncharacterized Zn-finger protein	COG01326													1		
arCOG05631	S	Function unknown	-	Uncharacterized protein														1		
arCOG05717	S	Function unknown	-	Uncharacterized protein	SC.00530													1		
arCOG05739	S	Function unknown	-	Uncharacterized protein with conserved CXXC pairs, DUF1667 family	COG03862	pfam07892												1		
arCOG05759	S	Function unknown	-	Uncharacterized protein		pfam12646												1		
arCOG05783	S	Function unknown	-	Uncharacterized protein														1		
arCOG05800	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG05803	S	Function unknown	-	Uncharacterized protein														1		
arCOG06053	S	Function unknown	-	Uncharacterized protein														1		
arCOG06113	S	Function unknown	-	Zn finger protein, C2C2 type	SC.00488													1		
arCOG06115	S	Function unknown	-	Pleckstrin homology domain containing proteins	SC.00612	pfam14470,pfam09851												1		
arCOG06521	S	Function unknown	-	Uncharacterized protein														1		
arCOG06692	S	Function unknown	-	Uncharacterized protein														1		
arCOG06939	S	Function unknown	-	WbqC-like protein family		pfam08889												1		
arCOG07169	S	Function unknown	-	Uncharacterized protein														1		
arCOG07353	S	Function unknown	-	Zn-ribbon protein														1		
arCOG07356	S	Function unknown	-	Uncharacterized protein														1		
arCOG07411	S	Function unknown	-	Uncharacterized membrane protein	SC.00711													1		
arCOG08157	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG08946	S	Function unknown	-	Uncharacterized protein														1		
arCOG09581	S	Function unknown	-	Uncharacterized membrane protein, DUF1700 family	COG04709	pfam08006												1		
arCOG09752	S	Function unknown	-	Uncharacterized membrane protein, DUF2068	COG03305													1		
arCOG10088	S	Function unknown	-	Uncharacterized protein	COG05504													1		
arCOG10150	S	Function unknown	-	Uncharacterized protein	SC.00215													1		
arCOG10658	S	Function unknown	-	Uncharacterized protein														1		
arCOG10671	S	Function unknown	-	Uncharacterized protein														1		
arCOG10940	S	Function unknown	-	Uncharacterized membrane protein	SC.00942													1		
arCOG11614	S	Function unknown	-	Uncharacterized protein														1		
arCOG11996	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13491	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13492	S	Function unknown	-	Uncharacterized protein														1		
arCOG13493	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13494	S	Function unknown	-	Uncharacterized protein														1		
arCOG13495	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13496	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13498	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13499	S	Function unknown	-	Uncharacterized protein, DUF4125 family		pfam13526												1		
arCOG13500	S	Function unknown	-	Uncharacterized protein														1		
arCOG13502	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13503	S	Function unknown	-	Uncharacterized protein														1		
arCOG13507	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13508	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13509	S	Function unknown	-	Uncharacterized protein														1		
arCOG13511	S	Function unknown	-	Uncharacterized protein														1		
arCOG13515	S	Function unknown	-	Uncharacterized protein														1		
arCOG13516	S	Function unknown	-	Uncharacterized protein														1		
arCOG13517	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13518	S	Function unknown	-	Uncharacterized protein														1		
arCOG13520	S	Function unknown	-	Uncharacterized protein		pfam02677												1		
arCOG13521	S	Function unknown	-	Uncharacterized membrane protein	SC.01004													1		
arCOG13523	S	Function unknown	-	Uncharacterized protein														1		
arCOG13524	S	Function unknown	-	Uncharacterized protein	SC.00667													1		
arCOG13525	S	Function unknown	-	Uncharacterized protein														1		
arCOG13526	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13527	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13528	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13529	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13530	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13533	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13534	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13535	S	Function unknown	-	Uncharacterized protein														1		
arCOG13539	S	Function unknown	-	Uncharacterized protein														1		
arCOG13540	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13541	S	Function unknown	-	Uncharacterized protein														1		
arCOG13542	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13543	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13544	S	Function unknown	-	Uncharacterized protein														1		
arCOG13545	S	Function unknown	-	Uncharacterized protein														1		
arCOG13547	S	Function unknown	-	Uncharacterized protein														1		
arCOG13548	S	Function unknown	-	Uncharacterized protein														1		
arCOG13549	S	Function unknown	-	Uncharacterized protein														1		
arCOG13551	S	Function unknown	-	Uncharacterized membrane protein														1		

arCOG13552	S	Function unknown	-	Uncharacterized protein														1		
arCOG13553	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13554	S	Function unknown	-	Uncharacterized protein														1		
arCOG13555	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13556	S	Function unknown	-	Uncharacterized protein														1		
arCOG13557	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13559	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13560	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13561	S	Function unknown	-	Uncharacterized protein														1		
arCOG13562	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13563	S	Function unknown	-	Uncharacterized protein														1		
arCOG13564	S	Function unknown	-	Uncharacterized membrane protein	SC.01004													1		
arCOG13565	S	Function unknown	-	Uncharacterized protein														1		
arCOG13566	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG13567	S	Function unknown	-	Uncharacterized protein														1		
arCOG15190	S	Function unknown	-	Uncharacterized membrane protein														1		
arCOG02545	S	Function unknown	-	Secreted protein, with PKD repeat domain	COG03291	pfam13229,pfam13229,pfam00801		TIGR04247										2		
arCOG03338	S	Function unknown	-	Uncharacterized membrane protein	SC.00033													2		
arCOG03817	S	Function unknown	-	Uncharacterized protein	SC.00100													2		
arCOG05407	S	Function unknown	-	Uncharacterized protein														2		
arCOG05805	S	Function unknown	-	Uncharacterized protein	SC.00141													2		
arCOG08126	S	Function unknown	-	Uncharacterized membrane protein	SC.00775													2		
arCOG09431	S	Function unknown	-	Uncharacterized protein		pfam14478												2		
arCOG10041	S	Function unknown	-	Uncharacterized membrane protein HdeD, DUF308 family	COG03247													2		
arCOG12007	S	Function unknown	-	Uncharacterized protein														2		
arCOG13538	S	Function unknown	-	Uncharacterized membrane protein	SC.00702													2		
arCOG02994	S	Function unknown	-	Cupin domain containing protein	COG01917	pfam07883												3		
arCOG10113	S	Function unknown	-	Uncharacterized protein														3		
arCOG10161	S	Function unknown	-	Uncharacterized protein														3		
arCOG03118	S	Function unknown	-	Membrane protein DegA family	COG01238	pfam09335													1	1
arCOG07655	S	Function unknown	-	Uncharacterized protein															1	1
arCOG01713	S	Function unknown	-	Uncharacterized protein	COG05440	pfam10061													1	3
arCOG02561	S	Function unknown	-	WD40 repeats containing protein	COG02319	pfam13191,pfam1160	cd00200,cd002	TIGR02794,TIGR02800,TIGR02800											1	3
arCOG00620	S	Function unknown	-	Uncharacterized conserved DUF39 domain fused to CBS domain	COG01900	pfam01837,pfam0057	cd04605	TIGR03287,TIGR01302											1	
arCOG04574	S	Function unknown	LemA	Uncharacterized protein	COG01704	pfam04011													1	
arCOG04811	S	Function unknown	-	Uncharacterized membrane protein, Fun14 family	COG02383														1	
arCOG04848	S	Function unknown	-	Uncharacterized protein, DUF1722 family	COG03272	pfam04463,pfam08349													1	
arCOG06436	S	Function unknown	-	Uncharacterized protein, YdiU/UPF0061 family	COG00397	pfam02696													1	
arCOG06646	S	Function unknown	-	Uncharacterized membrane protein, YccA/Bax inhibitor family	COG04760	pfam12811													1	
arCOG14289	S	Function unknown	-	Uncharacterized protein															1	
arCOG10710	S	Function unknown	-	Uncharacterized protein															2	
arCOG01908	S	Function unknown	AIM24	Uncharacterized protein, AIM24 family	COG02013	pfam01987														1
arCOG03873	S	Function unknown	-	Uncharacterized membrane protein	COG02855	pfam03601		TIGR00698												1
arCOG04987	S	Function unknown	-	Uncharacterized protein																1
arCOG05100	S	Function unknown	-	Uncharacterized membrane protein	COG02364															1
arCOG05340	S	Function unknown	-	Uncharacterized membrane protein																1
arCOG05874	S	Function unknown	-	Uncharacterized protein																1
arCOG06493	S	Function unknown	-	HEAT repeats containing protein	COG01413	pfam13646,pfam03130														1
arCOG10568	S	Function unknown	-	Uncharacterized membrane protein																1
arCOG02539	S	Function unknown	-	Secreted protein, with PKD repeat domain	COG03291															2
arCOG02614	S	Function unknown	-	Uncharacterized conserved protein																2
arCOG10301	S	Function unknown	-	Uncharacterized protein	COG03011	pfam04134														2
arCOG02284	R	General function prediction only	-	Secreted protein containing C-terminal beta-propeller domain distantly relate	COG04880	pfam09826			1	1	1		1	1	1	1			1	1
arCOG08640	R	General function prediction only	-	Predicted RNA-binding protein containing PUA-domain	COG02947	pfam01878			1	1	2			1	1	1			1	4
arCOG06896	R	General function prediction only	RRM	RNA-binding protein, RRM domain	COG00724	pfam00076	cd12399	TIGR01622	1	1		1			1	1				
arCOG00833	R	General function prediction only	RimI	Acetyltransferase (GNAT) family	COG00456	pfam00583	cd04301	TIGR01575	1	1					1	1		1	1	
arCOG02452	R	General function prediction only	-	KaiC-like protein ATPase	COG00467				1	1					1	1		3	1	
arCOG04430	R	General function prediction only	-	HD superfamily phosphohydrolase	COG01078	pfam01966	cd00077		1	2	1	1	1	1	1	1		1	1	2
arCOG02680	R	General function prediction only	-	Uncharacterized archaeal Zn-finger protein	COG01326				1	2	1	1	1	1	1	1		1	1	
arCOG04227	R	General function prediction only	-	Predicted CoA-binding protein	COG01832	pfam13380		TIGR02717	1	2	1	1			1	1			1	1
arCOG00557	R	General function prediction only	Lhr	Lhr-like helicase	COG01201	pfam00270,pfam0027	cd00046,cd000	TIGR04121	1		1	1	1		1	1		2	1	2
arCOG01150	R	General function prediction only	-	Predicted phosphohydrolase, MPP superfamily	COG01407		cd07391	TIGR00024	1		1	1	1		1	1		2		
arCOG01857	R	General function prediction only	-	SpoU rRNA Methylase family enzyme	COG01303	pfam01994			1		1	1		1	1	1		1	1	2
arCOG04151	R	General function prediction only	-	ABC-type multidrug transport system, permease component	COG02237	pfam04123			1		1	1		1	1	1		1	1	2
arCOG09205	R	General function prediction only	-	Predicted metal-dependent hydrolase	COG01547	pfam03745			1		1		1	1	1	1				
arCOG09415	R	General function prediction only	-	Membrane associated metal-binding domain fused to Reeler domain		pfam02014	cd08544		1		1			2	1	1			2	6
arCOG00626	R	General function prediction only	TlyC	Hemolysins or related protein containing CBS domains	COG01253	pfam01595,pfam0057	cd04590	TIGR03520	1			1			1	1			1	2
arCOG02291	R	General function prediction only	-	HAD superfamily hydrolase	COG01011	pfam13419	cd01427	TIGR02252							1	1		2		
arCOG00979	R	General function prediction only	-	Predicted O-methyltransferase YrrM	COG04122	pfam13578	cd02440								1	1				
arCOG01619	R	General function prediction only	ARA1	Aldo/keto reductase, related to diketogulonate reductase	COG00656	pfam00248	cd06660	TIGR01293	1	1				1	1	2			1	1
arCOG04807	R	General function prediction only	OPT	Uncharacterized membrane protein, oligopeptide transporter (OPT) family	COG01297	pfam03169		TIGR00733	2	3	1	1			1	2			2	1
arCOG00517	R	General function prediction only	-	Rhodanese Homology Domain fused to Zn-dependent hydrolase of beta-lacta	COG00491	pfam00753		TIGR03413	1	2	1				1	3			2	1
arCOG01360	R	General function prediction only	-	MiaB family, Radical SAM enzyme	COG01244	pfam04055	cd01335	TIGR01210	1	1	1	1		1	1			1	1	1
arCOG04278	R	General function prediction only	-	Predicted aconitase	COG01679	pfam04412	cd01355		1	1	1	1		1	1			1	1	1
arCOG00215	R	General function prediction only	CinA	Predicted nucleotide-utilizing enzyme related to molybdopterin-biosynthesis	COG01058	pfam00994	cd00885	TIGR00200	1	1	1	1		1	1					
arCOG00505	R	General function prediction only	-	Metal-dependent hydrolase of the beta-lactamase superfamily II	COG00491	pfam00753		TIGR03413	1	1	1	1		1	1					
arCOG01641	R	General function prediction only	-	Predicted RNA-binding protein, contains TRAM domain	COG03269				1	1	1	1			1			1		1
arCOG00504	R	General function prediction only	-	Metal-dependent hydrolase of the beta-lactamase superfamily II	COG00491	pfam00753		TIGR03413	1	1	1	1			1				3	1
arCOG00348	R	General function prediction only	-	Archaeal enzyme of ATP-grasp superfamily	COG02047	pfam09754		TIGR00162	1	1		1	1	1	1			1	1	2
arCOG04055	R	General function prediction only	-	SHS2 domain protein implicated in nucleic acid metabolism	COG01371	pfam01951			1	1		1			1			1	1	1
arCOG00313	R	General function prediction only	Sco1	Cytochrome oxidase Cu insertion factor, SCO1/SenC/PrrC family	COG01999	pfam02630	cd02968		1	1			1		1				2	

arCOG00893	R	General function prediction only	-	Predicted metal-dependent hydrolase (urease superfamily)	COG01831	pfam01026			1	1					1		1	1	2
arCOG00543	R	General function prediction only	-	Predicted metal-dependent RNase, consists of a metallo-beta-lactamase dom	COG01782	pfam00753,pfam1099	cd02410	TIGR03675	1		1	1	1	1	1		1	1	5
arCOG01093	R	General function prediction only	-	Protein distantly related to bacterial ferritins	COG02406	pfam00210	cd01052		1			1			1				
arCOG00347	R	General function prediction only	-	Archaeal enzyme of ATP-grasp superfamily	COG01938	pfam09754		TIGR00161	2		1	1	1	1	1		1	2	1
arCOG04628	R	General function prediction only	-	Predicted metal-dependent hydrolase	COG01547	pfam03745				1	1	1			1				
arCOG02579	R	General function prediction only	-	Fe-S-cluster containining protein	COG00727					1					1			2	1
arCOG03096	R	General function prediction only	-	NAD dependent epimerase/dehydratase family enzyme	COG01090	pfam01370,pfam0833	cd05242	TIGR01777		2				1	1			1	3
arCOG03991	R	General function prediction only	-	PKD repeats containing protein	COG03291	pfam00801	cd00146	TIGR00864,TIGR04213					1		1				2
arCOG04212	R	General function prediction only	-	Predicted DNA-binding protein with PD1-like DNA-binding motif	COG01661	pfam03479	cd11378								1		1	1	1
arCOG06747	R	General function prediction only	-	SIR2 superfamily protein	COG00846	pfam02146,pfam1328	cd01406								1				
arCOG06769	R	General function prediction only	-	GTPase SAR1 family domain fused to Leucine-rich repeats domain	COG01100	pfam13855,pfam1385	cd00116,cd099	TIGR00231							1				
arCOG00499	R	General function prediction only		Metal-dependent hydrolase of the beta-lactamase superfamily	COG01235	pfam12706		TIGR02651	1	1	1	1	1	1	2	1	1	1	1
arCOG01225	R	General function prediction only	-	GTPase SAR1 or related small G protein	COG01100	pfam03029	cd02027,cd00880		1	1	1	1	1	1	2	1	1	1	
arCOG04444	R	General function prediction only	-	ACT domain containing protein	COG04747	pfam01842,pfam0184	cd04908,cd04908		1	1	1	1	1		2	1			2
arCOG00062	R	General function prediction only	-	Predicted amidohydrolase	COG00388	pfam00795	cd07197	TIGR03381	1	2	1	1	1	1	2	1	2		
arCOG02293	R	General function prediction only	-	HAD superfamily hydrolase	COG00637	pfam13419	cd01427	TIGR02009	2	3	1	2	1	2	2	1	2		
arCOG02742	R	General function prediction only	-	Uncharacterized membrane anchored protein with extracellular vWF domain	COG01721	pfam01882			2	1	1	2			2	2	1	2	4
arCOG04303	R	General function prediction only	-	Uncharacterized Rossmann fold enzyme	COG01634	pfam01973	cd07995		1	1	1		1		2		1	1	1
arCOG03450	R	General function prediction only	-	Predicted surface protease of transglutaminase family					1	2	1	1		1	2		1		
arCOG04466	R	General function prediction only	-	Na+-dependent transporter of the SNF family	COG00733	pfam00209	cd10336		1	2		1		1	2			1	3
arCOG04950	R	General function prediction only	YggS	Uncharacterized pyridoxal phosphate-containing protein, affects Ilv metabolis	COG00325	pfam01168	cd00635	TIGR00044	1						2		1		
arCOG04624	R	General function prediction only	-	Multimeric flavodoxin WrbA	COG00431				1						2			1	1
arCOG02155	R	General function prediction only	-	Protein implicated in RNA metabolism, contains PRC-barrel domain	COG01873	pfam05239			1	1	1	1	1	1	4	1	2	1	2
arCOG03682	R	General function prediction only	SPS1	Membrane associated serine/threonine protein kinase	COG00515	pfam00069	cd14014		2	3	2	1	2		4	2		1	
arCOG02174	R	General function prediction only	-	Predicted exporter of the RND superfamily	COG01033	pfam03176,pfam03176		TIGR00921	4	4	4	3		3	5	1	1	6	11
arCOG01848	R	General function prediction only	WbbJ	Acetyltransferase (isoleucine patch superfamily)	COG00110	pfam00132,pfam0013	cd03358	TIGR03570	1	2				1		1		1	1
arCOG06534	R	General function prediction only	-	Cohesin domain containing secreted protein		pfam00963	cd08547		1			1				1		1	
arCOG07590	R	General function prediction only	-	Predicted acyl esterase	COG02936	pfam02129,pfam0853	0	TIGR00976	3	4	1	2	2	2		1		2	5
arCOG05343	R	General function prediction only	-	GTPase SAR1 or related small G protein	COG01100	pfam00071	cd00154	TIGR00231				1		1		1	1		
arCOG03439	R	General function prediction only	-	vWFa domain containing protein								2				1	1	2	3
arCOG04179	R	General function prediction only	PDCD5	DNA-binding TFAR19-related protein, PDSD5 family	COG02118	pfam01984										1	1	1	
arCOG00601	R	General function prediction only	-	Protein containing two CBS domains (some fused to C-terminal double-strand	COG00517	pfam00571,pfam0057	cd04633,cd046	TIGR01302,TIGR01302								1	1		
arCOG02935	R	General function prediction only	-	Pirin-related protein	COG01741	pfam02678,pfam05726										1		1	
arCOG02289	R	General function prediction only	-	Uncharacterized protein, DUF169 family	COG02043	pfam02596										1			
arCOG03994	R	General function prediction only	-	PKD domain containing protein	COG03291	pfam05048,pfam0080	cd14251,cd001	TIGR04247,TIGR04275,TIGR04275,TIGR04275,TIGR04275,TIGR04275,TIGR04275,TIGR04								1			
arCOG06897	R	General function prediction only	-	Predicted ATP-grasp enzyme	COG03919	pfam15632		TIGR01369	3	3	1		1	1		4			
arCOG03038	R	General function prediction only	-	TPR repeats containing protein	COG00457				1	2	1	2	1	1		5	4	2	3
arCOG01285	R	General function prediction only	-	OB-fold domain and Zn-ribbon containing protein, possible acyl-CoA-binding p	COG01545	pfam12172,pfam0179	6		1	1	1	1					1		
arCOG01651	R	General function prediction only	-	Alpha/beta superfamily hydrolase	COG01073	pfam08538			1	1	1	2					2		2
arCOG02812	R	General function prediction only	-	Bacteriorhodopsin	COG05524	pfam01036			1	1	1								1
arCOG00352	R	General function prediction only	Nog1	GTP-binding protein, GTP1/Obg family	COG01084	pfam01926	cd01897	TIGR02729	1	1		1					1	1	1
arCOG01108	R	General function prediction only	AbgB	Metal-dependent amidase/aminoacylase/carboxypeptidase	COG01473	pfam01546	cd03886	TIGR01891	1	1		1					1		1
arCOG07416	R	General function prediction only	MhpC	Alpha/beta superfamily hydrolase	COG00596	pfam12695		TIGR03695	1	1								1	1
arCOG05195	R	General function prediction only	-	TPR repeats containing protein	COG00457	pfam13414	cd00189	TIGR02521	1	1									2
arCOG02889	R	General function prediction only	-	Predicted deacylase	COG03608	pfam04952	cd06252	TIGR02994	1	1									
arCOG01425	R	General function prediction only	-	RecB family nuclease with coiled-coil N-terminal domain	SC.00001				1		1	1	1						1
arCOG02259	R	General function prediction only	-	Zn-finger domain containing protein	SC.00304				1			1	1	1					
arCOG01234	R	General function prediction only	-	GTPase, G3E family	COG00523	pfam02492,pfam0768	cd03112	TIGR02475	1			1							2
arCOG01141	R	General function prediction only	-	ICC-like phosphoesterase	COG00622	pfam12850	cd00841	TIGR00040	1								1		
arCOG01395	R	General function prediction only	-	Lipid-A-disaccharide synthase related glycosyltransferase	COG01817	pfam04007			1										
arCOG01648	R	General function prediction only	MhpC	Alpha/beta superfamily hydrolase	COG00596	pfam12697		TIGR02427	1										
arCOG02164	R	General function prediction only	-	Predicted transglutaminase-like protease	COG01800	pfam04473			1										
arCOG01622	R	General function prediction only	MviM	Predicted dehydrogenase	COG00673	pfam01408,pfam0289	4		2					1				1	
arCOG03509	R	General function prediction only	-	Cell surface protein, containing PKD repeats	COG03291	pfam10102				1	1								
arCOG00614	R	General function prediction only	SpoIVFB	Zn-dependent protease	COG01994		cd06158			1							1	1	2
arCOG03015	R	General function prediction only	-	Uncharacterized protein YbjT, contains NAD(P)-binding and DUF2867 domains	COG00702	pfam13460	cd05271	TIGR03466		1									1
arCOG02625	R	General function prediction only	-	Predicted metal-dependent hydrolase	COG01451	pfam01863				1									
arCOG06095	R	General function prediction only	-	ABC-2 family transporter protein	COG01277	pfam12679				1									
arCOG01136	R	General function prediction only	ARC1	EMAP domain RNA-binding protein	COG00073	pfam01588	cd02800	TIGR00399		2									1
arCOG03048	R	General function prediction only	-	TPR repeats containing protein	COG00457	pfam13414,pfam1343	cd00189,cd001	TIGR02917			1	1	1						
arCOG04932	R	General function prediction only	-	Cell surface protein	COG01470						1			1					
arCOG01213	R	General function prediction only	Cof	HAD superfamily hydrolase	COG00561	pfam08282	cd01427	TIGR01487			1							1	2
arCOG03850	R	General function prediction only	-	Predicted O-methyltransferase YrrM	COG04122	pfam13578								1					
arCOG00302	R	General function prediction only	-	Metal-dependent phosphoesterase (PHP family)	COG00613	pfam02811,pfam1326	cd07432										1	1	1
arCOG00353	R	General function prediction only	HflX	GTP-binding protein protease modulator	COG02262	pfam13167,pfam0192	cd01878	TIGR03156									1	1	1
arCOG01043	R	General function prediction only	-	Predicted RNA binding protein with dsRBD fold	COG01931	pfam01877											1	1	1
arCOG04589	R	General function prediction only	CinA	Uncharacterized protein (competence- and mitomycin-induced)	COG01546	pfam02464		TIGR00199									1	1	1
arCOG05137	R	General function prediction only	-	TPR repeats containing protein	COG00457	pfam13414,pfam1341	cd00189,cd001	TIGR02917									1	1	1
arCOG00932	R	General function prediction only	-	Uncharacterized protein related to pyruvate formate-lyase activating enzyme	COG02108	pfam04055	cd01335										1	1	2
arCOG04116	R	General function prediction only	-	ATPase (PilT family)	COG01855	pfam00437	cd09878,cd01130										1	1	2
arCOG04265	R	General function prediction only	-	C4-type Zn-finger protein	COG01779	pfam03367		TIGR00340									1	1	3
arCOG01084	R	General function prediction only	MazG	Predicted pyrophosphatase	COG01694	pfam03819	cd11535										1	1	
arCOG02448	R	General function prediction only	-	Uncharacterized Fe-S center protein	COG02768	pfam04015,pfam12838		TIGR01971									1		2
arCOG00254	R	General function prediction only	-	Predicted dinucleotide-utilizing enzyme	COG01712	pfam03447,pfam01958		TIGR03855									1		
arCOG00600	R	General function prediction only	-	CBS domain	COG00517	pfam00571,pfam0057	cd04623,cd046	TIGR01302,TIGR00393									1		
arCOG00662	R	General function prediction only	-	Biotin synthase-related protein, radical SAM superfamily	COG02516	pfam04055	cd01335	TIGR04317,TIGR04043									1		
arCOG00691	R	General function prediction only	-	Predicted metal-dependent hydrolase with the TIM-barrel fold	COG01574	pfam07969	cd01300										1		
arCOG00769	R	General function prediction only	ThiJ	Putative intracellular protease/amidase	COG00693	pfam01965	cd03134	TIGR01382									1		
arCOG00828	R	General function prediction only	-	Acetyltransferase GNAT superfamily	COG03393	pfam08445		TIGR01575									1		
arCOG00913	R	General function prediction only	-	Predicted methyltransferase	COG01568	pfam01861											1		
arCOG00933	R	General function prediction only	-	Radical SAM superfamily enzyme	COG01964	pfam04055	cd01335	TIGR02666									1		

arCOG00940	R	General function prediction only	-	Radical SAM superfamily enzyme	COG00535	pfam04055,pfam1318	cd01335	TIGR04055										1		
arCOG00969	R	General function prediction only	-	Predicted hydrolase (metallo-beta-lactamase superfamily)	COG02248	pfam12706												1		
arCOG01145	R	General function prediction only	-	Phosphohydrolase, Icc/MPP superfamily	COG02129	pfam00149	cd07392											1		
arCOG01263	R	General function prediction only	DltE	Short-chain dehydrogenase	COG00300		cd05233	TIGR01830										1		
arCOG01294	R	General function prediction only	HcaD	NAD(FAD)-dependent dehydrogenase	COG00446	pfam13510,pfam0799	cd00207,cd052	TIGR01372										1		
arCOG01295	R	General function prediction only	HcaD	NAD(FAD)-dependent dehydrogenase	COG00446	pfam07992		TIGR01292										1		
arCOG01377	R	General function prediction only	-	Phosphodiesterase of AP superfamily	COG03379	pfam01663												1		
arCOG01378	R	General function prediction only	-	Uncharacterized protein of the AP superfamily	COG01524	pfam01663												1		
arCOG01455	R	General function prediction only	AdhP	Zn-dependent alcohol dehydrogenase	COG01064	pfam08240,pfam0010	cd08259	TIGR02824										1		
arCOG01492	R	General function prediction only	YjgC	Predicted molibdopterin-dependent oxidoreductase YjgC	COG03383	pfam13510,pfam1058	cd00207,cd027	TIGR02512,TIGR01591										1		
arCOG01728	R	General function prediction only	Mho1	Predicted class III extradiol dioxygenase, MEMO1 family	COG01355	pfam01875	cd07361	TIGR04336										1		
arCOG01801	R	General function prediction only	Imp	TRAP-type uncharacterized transport system, periplasmic component	COG02358	pfam12974	cd13567	TIGR02122										1		
arCOG01849	R	General function prediction only	PaaY	Isoleucine patch superfamily protein	COG00663		cd04650	TIGR02287										1		
arCOG01906	R	General function prediction only	-	TRAP-type uncharacterized transport system, fused permease component	COG04666	pfam06808		TIGR02123										1		
arCOG02008	R	General function prediction only	-	Uncharacterized membrane protein, a putative transporter component	COG03371	pfam06197												1		
arCOG02238	R	General function prediction only	-	Predicted DNA-binding protein	COG01342	pfam02001		TIGR02937										1		
arCOG02317	R	General function prediction only	-	Predicted regulator of amino acid metabolism, contains ACT domain	COG02150													1		
arCOG02431	R	General function prediction only	-	Predicted Rossmann fold nucleotide-binding protein	COG01611			TIGR00725										1		
arCOG02444	R	General function prediction only	-	Predicted permease	COG03368	pfam09847												1		
arCOG02828	R	General function prediction only	-	NAD dependent epimerase/dehydratase family	COG03367	pfam07755												1		
arCOG02900	R	General function prediction only	-	vWFA domain containing protein	COG02304	pfam00092	cd00198											1		
arCOG02902	R	General function prediction only	-	vWFA domain containing protein	COG02304	pfam13519	cd01467											1		
arCOG03045	R	General function prediction only	-	TPR repeats containing protein	COG00457	pfam13424	cd00189,cd00189											1		
arCOG03427	R	General function prediction only	-	DMT superfamily transporter	COG02510													1		
arCOG03639	R	General function prediction only	-	Predicted glutamine amidotransferase	COG00121		cd01908											1		
arCOG03691	R	General function prediction only	-	Cell surface protein														1		
arCOG04065	R	General function prediction only	PqqL	Zn-dependent peptidase	COG00612	pfam00675,pfam05193												1		
arCOG04115	R	General function prediction only	SfsA	DNA-binding protein, stimulates sugar fermentation	COG01489	pfam03749		TIGR00230										1		
arCOG04230	R	General function prediction only	-	HD supefamily hydrolase	COG03294													1		
arCOG04290	R	General function prediction only	-	PIN-domain and Zn ribbon	COG01656	pfam01927												1		
arCOG04291	R	General function prediction only	-	Sugar isomerase related protein	COG01801	pfam01904												1		
arCOG04331	R	General function prediction only	-	Thioesterase-like protein	COG05496		cd03440											1		
arCOG04354	R	General function prediction only	-	Tripartite tricarboxylate transporter (TTT) class transporter	COG01906	pfam04165		TIGR00529										1		
arCOG04359	R	General function prediction only	-	Predicted RNA-binding protein containing a C-terminal EMAP domain	COG02517	pfam01588	cd02796	TIGR00472										1		
arCOG04410	R	General function prediction only	-	Predicted ATP-grasp domain fused to redox center	COG01578	pfam01937												1		
arCOG04418	R	General function prediction only	-	Predicted HTH domain, homologous to N-terminal domain of RPA1 protein fa	COG03612	pfam09999												1		
arCOG04426	R	General function prediction only	HybF	Zn finger protein HypA/HybF (possibly regulating hydrogenase expression)	COG00375	pfam01155		TIGR00100										1		
arCOG04477	R	General function prediction only	-	Predicted metal binding protein, contains two cysteine clusters	COG01860	pfam03684												1		
arCOG05097	R	General function prediction only	-	CBS domain	COG00517	pfam00571,pfam0057	cd02205	TIGR01302										1		
arCOG05366	R	General function prediction only	-	Cell surface protein	COG01572	pfam07705												1		
arCOG05810	R	General function prediction only	-	Uncharacterized protein, homolog of lactam utilization protein B	COG01540	pfam03746	cd10787											1		
arCOG05825	R	General function prediction only	-	Radical SAM superfamily enzyme	COG01856	pfam04055	cd01335											1		
arCOG07368	R	General function prediction only	-	Cell surface protein	COG01572	pfam07705												1		
arCOG11082	R	General function prediction only	-	Uncharacterized membrane protein, a component of a putative secretion system														1		
arCOG00497	R	General function prediction only	-	Zn-dependent hydrolase of the beta-lactamase fold	COG02220	pfam13483												2	1	
arCOG00606	R	General function prediction only	-	CBS domain	COG00517	pfam00478	cd04623	TIGR01302										2	1	
arCOG04469	R	General function prediction only	-	Tripartite tricarboxylate transporter (TTT) class transporter	COG01784	pfam01970												2	1	
arCOG03032	R	General function prediction only	-	TPR repeats containing protein	COG00457	pfam13414,pfam1341	cd00189	TIGR02917										2		1
arCOG00503	R	General function prediction only	-	Metal-dependent hydrolase of the beta-lactamase superfamily II	COG01237	pfam00753		TIGR03675										2		
arCOG01688	R	General function prediction only	-	Predicted transcriptional regulator, contains HTH and 4VR domain	COG01719	pfam02830												2		
arCOG01963	R	General function prediction only	-	PhoU-like domain fused to TrkA-C domain	COG03273	pfam01895,pfam02080												2		
arCOG02292	R	General function prediction only	Gph	HAD superfamily hydrolase	COG00546	pfam13419	cd01427	TIGR03351										2		
arCOG02603	R	General function prediction only	-	Roadblock/LC7 domain	COG02018													2		
arCOG03167	R	General function prediction only	-	Predicted ATPase, AAA+ superfamily	COG01373	pfam13173,pfam13635												2		
arCOG03400	R	General function prediction only	-	Predicted RNA-binding protein, contains TRAM domain	COG04085	pfam01336	cd04485											2		
arCOG04409	R	General function prediction only	-	Predicted nuclease (RNAse H fold)	COG02410	pfam04250												2		
arCOG07997	R	General function prediction only	-	Predicted OB fold RNA-binding domain fused metal-dependent hydrolase	COG01988	pfam01336,pfam04307												2		
arCOG00938	R	General function prediction only	-	Radical SAM superfamily enzyme	COG00535	pfam04055,pfam1318	cd01335	TIGR04084										5		
arCOG02175	R	General function prediction only	-	Predicted transporter of the RND superfamily	COG02409	pfam03176,pfam03176		TIGR00833,TIGR00833											1	1
arCOG10597	R	General function prediction only	-	Hemocyanin family protein, binds copper ions		pfam00264													1	1
arCOG11383	R	General function prediction only	-	MOCS domain, sulfur-carrier protein	COG02258	pfam03473													1	1
arCOG00498	R	General function prediction only	-	Metal-dependent hydrolase of the beta-lactamase superfamily II	COG00491	pfam00753		TIGR03413											1	2
arCOG00500	R	General function prediction only	ElaC	Metal-dependent hydrolase of the beta-lactamase superfamily	COG01234	pfam12706		TIGR02651											1	2
arCOG00502	R	General function prediction only	PhnP	Metal-dependent hydrolase of the beta-lactamase superfamily	COG01235														1	2
arCOG02642	R	General function prediction only	PerM	Predicted PurR-regulated permease PerM	COG00628	pfam01594		TIGR02872											1	2
arCOG02890	R	General function prediction only	-	Predicted deacylase	COG03608	pfam04952	cd06251	TIGR02994											1	2
arCOG04743	R	General function prediction only	-	Helicase associated uncharacterized terminal domain		pfam05854													1	2
arCOG00654	R	General function prediction only	-	Predicted periplasmic solute-binding protein	COG02107	pfam02621	cd13534												1	3
arCOG00655	R	General function prediction only	-	Predicted periplasmic solute-binding protein	COG01427	pfam02621	cd13634												1	3
arCOG01189	R	General function prediction only	AarF	Predicted unusual protein kinase	COG00661	pfam03109	cd05121	TIGR01982											1	3
arCOG01030	R	General function prediction only	-	Predicted kinase related to galactokinase and mevalonate kinase	COG02605														1	6
arCOG01850	R	General function prediction only	WbbJ	Acetyltransferase (isoleucine patch superfamily)	COG00110	pfam00132	cd04647	TIGR03532											1	
arCOG02303	R	General function prediction only	SurE	Predicted acid phosphatase	COG00496	pfam01975,pfam14423		TIGR00087											1	
arCOG02839	R	General function prediction only	-	Uncharacterized protein related to deoxyribodipyrimidine photolyase	COG03046	pfam04244,pfam03441													1	
arCOG02986	R	General function prediction only	BioY	Uncharacterized protein	COG01268	pfam02632													1	
arCOG03047	R	General function prediction only	-	TPR repeats containing protein	COG00457														1	
arCOG03241	R	General function prediction only	-	ATPase AAA family	COG04637	pfam13304,pfam13304													1	
arCOG03271	R	General function prediction only	YtfP	Uncharacterized protein YtfP, gamma-glutamylcyclotransferase (GGCT)/AIG2-	COG02105	pfam06094	cd06661												1	
arCOG05099	R	General function prediction only	YtfP	Uncharacterized protein YtfP, gamma-glutamylcyclotransferase (GGCT)/AIG2-	COG02105	pfam06094	cd06661												1	
arCOG08119	R	General function prediction only	PhoX	Secreted phosphatase, PhoX family	COG03211	pfam05787													1	
arCOG03561	R	General function prediction only	-	Secreted protein with beta-propeller repeat domain	COG03391	pfam01436,pfam0143	cd14955												2	

arCOG07781	R	General function prediction only	-	Cell surface protein														2		
arCOG02562	R	General function prediction only	-	Beta-propeller repeat containing protein	COG03391													3		1
arCOG02810	R	General function prediction only	-	Bacteriorhodopsin	COG05524	pfam01036														1
arCOG03169	R	General function prediction only	-	AAA+ superfamily ATPase	COG01672	pfam01637,pfam01978														1
arCOG07790	R	General function prediction only	-	D-glucuronyl C5-epimerase C-terminal domain related protein		pfam06662														1
arCOG02316	R	General function prediction only	-	Predicted regulator of amino acid metabolism, contains ACT domain	COG02150															2
arCOG02560	R	General function prediction only	-	Secreted protein with beta-propeller repeat domain	COG03391		cd05819	TIGR03866												2
arCOG06256	R	General function prediction only	-	Predicted esterase	COG00400	pfam02230														2
arCOG00082	EF		PucG	Serine-pyruvate aminotransferase/archaeal aspartate aminotransferase	COG00075	pfam00266	cd06451	TIGR03301	1								1	1	1
arCOG01446	VK		Csa3	CRISPR-Cas assicated transcriptional regulator, contains CARF and HTH domai	COG00640		cd09655	TIGR01884									1		

Supplementary Table 6. Classification of the unique genes based on the arCOG classification.

									Number of genes				
		arCOG clasification	gene	product	COG classification	pFAM domain	cdc	TIGR classification	EPIPELAGIC MGIII	BATHY1	BATHY2	Thalassoarchaea	MG2-GG3
INFORMATION STORAGE AND PROCESSING													
arCOG00415	L	Replication, recombination and repair	RecA	RecA/RadA recombinase	COG00468	pfam14520,pfa	cd01123	TIGR02236	1				
arCOG00439	L	Replication, recombination and repair	MCM2	Predicted ATPase involved in replication control, Cdc46/Mcm family	COG01241	pfam14551,pfa	cd00009		1				
arCOG04110	L	Replication, recombination and repair	PRI1	Eukaryotic-type DNA primase, catalytic (small) subunit	COG01467	pfam01896	cd04860	TIGR00335	1			1	1
arCOG00787	L	Replication, recombination and repair	-	UvrD/Rep family helicase fused to exonuclease family domain	COG02887	pfam12705	cd09637	TIGR01249,TIGR00372	1				
arCOG05109	L	Replication, recombination and repair	DnaQ	DNA polymerase III, epsilon subunit or related 3'-5' exonuclease	COG00847	pfam00929	cd06127	TIGR00573	1				1
arCOG01898	L	Replication, recombination and repair	Uve	UV damage repair endonuclease	COG04294	pfam03851		TIGR00629	1				
arCOG00469	L	Replication, recombination and repair	HolB	ATPase involved in DNA replication HolB, small subunit	COG00470	pfam13177,pfa	cd00009	TIGR02397	2				1
arCOG01894	L	Replication, recombination and repair	Nfo	Endonuclease IV	COG00648	pfam01261	cd00019	TIGR00587	2				
arCOG01073	L	Replication, recombination and repair	-	NUDIX family hydrolase	COG00494	pfam00293	cd03424	TIGR00052	2				
arCOG01347	L	Replication, recombination and repair	CDC9	ATP-dependent DNA ligase	COG01793	pfam04675,pfa	cd07901,cd079	TIGR00574	2				
arCOG02840	L	Replication, recombination and repair	PhrB	Deoxyribodipyrimidine photolyase	COG00415	pfam00875,pfam03441		TIGR03556	2				1
arCOG01078	L	Replication, recombination and repair	-	NUDIX family hydrolase	COG00494	pfam00293	cd03428		3	1		1	1
arCOG00464	L	Replication, recombination and repair	AlkA	3-methyladenine DNA glycosylase/8-oxoguanine DNA glycosylase	COG00122	pfam07934,pfa	cd00056	TIGR00588	3		1		1
arCOG03142	L	Replication, recombination and repair	-	Nuclease of RNase H fold, RuvC/YqgF family					3		1	1	1
arCOG01486	L	Replication, recombination and repair	RnmV	5S rRNA maturation endonuclease (Ribonuclease M5), contains TOPRIM dom	COG01658	pfam01751	cd01027		4	1		1	1
arCOG08649	L	Replication, recombination and repair	-	Topoisomerase IB	COG03569	pfam02919,pfa	cd00660,cd00659		4				1
arCOG00417	L	Replication, recombination and repair	RecA	RecA/RadA recombinase	COG00468		cd01394	TIGR02237	5				
arCOG04367	L	Replication, recombination and repair	GyrA	Type IIA topoisomerase (DNA gyrase/topo II, topoisomerase IV), A subunit	COG00188	pfam00521,pfa	cd00187	TIGR01063				2	
arCOG00872	L	Replication, recombination and repair	MPH1	ERCC4-like helicase	COG01111	pfam00270,pfa	cd00046,cd120	TIGR00643,TIGR00580,TIGR00596			1	2	1
arCOG00551	L	Replication, recombination and repair	-	DNA replication initiation complex subunit, GINS15 family	COG01711		cd11714			1			
arCOG00427	L	Replication, recombination and repair	RecJ/CdcS	Single-stranded DNA-specific exonuclease RecJ	COG00608					1	1		
arCOG02258	L	Replication, recombination and repair	-	RPA family protein, a subunit of RPA complex in P.furiosus	COG03390					1		2	
arCOG01166	L	Replication, recombination and repair	MutL	DNA mismatch repair enzyme (predicted ATPase)	COG00323	pfam13589,pfa	cd00075,cd007	TIGR00585				1	
arCOG04121	L	Replication, recombination and repair	RnhB	Ribonuclease HII	COG00164	pfam01351	cd07180	TIGR00729				2	
arCOG00470	L	Replication, recombination and repair	HolB	ATPase involved in DNA replication HolB, large subunit	COG00470	pfam00004	cd00009	TIGR02397				1	1
arCOG04371	L	Replication, recombination and repair	GyrB	Type IIA topoisomerase (DNA gyrase/topo II, topoisomerase IV), B subunit	COG00187	pfam02518,pfa	cd00075,cd008	TIGR01059				1	1
arCOG00488	L	Replication, recombination and repair	DnaN	DNA polymerase sliding clamp subunit (PCNA homolog)	COG00592	pfam00705,pfa	cd00577	TIGR00590				2	1
arCOG01527	L	Replication, recombination and repair	TopA	Topoisomerase IA	COG00550	pfam01751,pfa	cd03362,cd001	TIGR01057		1		1	
arCOG04050	L	Replication, recombination and repair	FEN1	5'-3' exonuclease	COG00258	pfam00752,pfa	cd09867	TIGR03674		1			
arCOG00558	L	Replication, recombination and repair	SrmB	Superfamily II DNA and RNA helicase	COG00513	pfam00270,pfa	cd00268,cd000	TIGR01389				1	2
arCOG00328	L	Replication, recombination and repair	PolB3	DNA polymerase PolB3	COG00417	pfam03104,pfa	cd05781,cd055	TIGR00592				4	1
arCOG01510	L	Replication, recombination and repair	RPA1	Single-stranded DNA-binding replication protein A (RPA), large (70 kD) subun	COG01599		cd04491,cd04491			2		2	1
arCOG01072	L	Replication, recombination and repair	-	NUDIX family hydrolase	COG00494	pfam00293	cd03426			1			
arCOG03013	L	Replication, recombination and repair	PRI2	Eukaryotic-type DNA primase, large subunit	COG02219	pfam04104	cd06560			1			1
arCOG00553	L	Replication, recombination and repair	BRR2	Replicative superfamily II helicase	COG01204	pfam00270,pfa	cd00046,cd000	TIGR04121				3	1
arCOG01526	L	Replication, recombination and repair	TopG2	Reverse gyrase	COG01110	pfam00270,pfa	cd00046,cd033	TIGR01054				2	
arCOG00368	L	Replication, recombination and repair	SbcC	ATPase involved in DNA repair, SbcC	COG00419	pfam13476,pfa	cd03240,cd001	TIGR00611,TIGR02168			1		
arCOG02724	L	Replication, recombination and repair	Ada	Methylated DNA-protein cysteine methyltransferase	COG00350	pfam01035	cd06445	TIGR00589			1	1	
arCOG00397	L	Replication, recombination and repair	SbcD	DNA repair exonuclease, SbcD	COG00420	pfam00149	cd00840				1		
arCOG00329	L	Replication, recombination and repair	PolB2	DNA polymerase PolB2, inactivated	COG00417	pfam00136	cd05531	TIGR00592				1	1
arCOG00802	L	Replication, recombination and repair	RecB	ATP-dependent exoDNAse (exonuclease V) beta subunit (contains helicase a	COG01074	pfam00580,pfam13361		TIGR02785					1
arCOG02895	L	Replication, recombination and repair	MutS2	DNA structure-specific ATPase involved in suppression of recombination, Mu	COG01193	pfam00488	cd03243	TIGR01069				1	1
arCOG04694	L	Replication, recombination and repair	UvrA	Excinuclease ABC subunit A, ATPase	COG00178		cd03271,cd032	TIGR00630				3	
arCOG00462	L	Replication, recombination and repair	MutY	A/G-specific DNA glycosylase	COG01194	pfam00730	cd00056	TIGR01084					1
arCOG03646	L	Replication, recombination and repair	XseB	Exonuclease VII small subunit	COG01722	pfam02609		TIGR01280					1
arCOG04513	L	Replication, recombination and repair	XseA	Exonuclease VII, large subunit	COG01570	pfam13742,pfa	cd04489	TIGR00237					1
arCOG04754	L	Replication, recombination and repair	Lig	NAD-dependent DNA ligase	COG00272	pfam01653,pfa	cd00114,cd000	TIGR00575					1
arCOG07300	L	Replication, recombination and repair	-	Uncharacterized protein associated with inactivated PolB-like polymerase									1
arCOG01082	L	Replication, recombination and repair	-	NUDIX family hydrolase	COG00494	pfam00293	cd02885					1	
arCOG01305	L	Replication, recombination and repair	-	Topoisomerase DNA binding C4 zinc finger fused to uncharacterized N-termi	COG01637							1	
arCOG02896	L	Replication, recombination and repair	MutS	Mismatch repair ATPase (MutS family)	COG00249	pfam01624,pfa	cd03284	TIGR01070				1	
arCOG01981	K	Transcription	SUA7	Transcription initiation factor TFIIB, Brf1 subunit/Transcription initiation fact	COG01405	pfam08271,pfa	cd00043		5	1		1	
arCOG01863	K	Transcription	-	Predicted transcription factor, homolog of eukaryotic MBF1	COG01813	pfam01381	cd00093	TIGR00270	3	1	1	1	
arCOG02611	K	Transcription	-	Predicted transcriptional regulator, containd two HTH domains	COG03398	pfam13412,pfa	cd00090,cd00090		3	1	1	2	1
arCOG07561	K	Transcription	-	Predicted membrane-associated trancriptional regulator	COG02512				3	1	1	2	4
arCOG00770	K	Transcription	DinG	Rad3-related DNA helicase	COG01199	pfam06733,pfam13307		TIGR00604	2		1	2	1
arCOG05671	K	Transcription	CsgD	Transcriptional regulator LuxR family	COG02771				2				
arCOG01680	K	Transcription	ArsR	Transcriptional regulator containing HTH domain, ArsR family	COG00640	pfam01022	cd00090		1		1		
arCOG04280	K	Transcription	NagC	Transcriptional regulator/sugar kinase	COG01940	pfam00480	cd00012	TIGR00744			1		
arCOG01753	K	Transcription	Ssh10b	Archaeal DNA-binding protein	COG01581	pfam01918		TIGR00285		1	1		
arCOG00579	K	Transcription	RPB9	DNA-directed RNA polymerase, subunit M/Transcription elongation factor TF	COG01594			TIGR01384				1	
arCOG01684	K	Transcription	-	Transcriptional regulator, ArsR family	COG01777	pfam01022	cd00090,cd00090					3	1
arCOG00675	K	Transcription	RPB7	DNA-directed RNA polymerase, subunit E'/Rpb7	COG01095	pfam03876,pfa	cd04331,cd044	TIGR00448					1
arCOG04258	K	Transcription	RPB5	DNA-directed RNA polymerase, subunit H, RpoH/RPB5	COG02012	pfam01191						1	1
arCOG01580	K	Transcription	-	DNA-binding transcriptional regulator, Lrp family	COG01522	pfam13412,pfa	cd00090			1		1	
arCOG05161	K	Transcription	WecD	Acetyltransferase (GNAT) family	COG00454					1			
arCOG04111	K	Transcription	RPB11	DNA-directed RNA polymerase, subunit L	COG01761	pfam13656	cd06927						1
arCOG01764	K	Transcription	SPT15	TATA-box binding protein (TBP), component of TFIID and TFIIIB	COG02101	pfam00352,pfa	cd04518			1			
arCOG02099	K	Transcription	TroR	Mn-dependent transcriptional regulator (DtxR family)	COG01321	pfam01325,pfam02742,pfam04023						1	1
arCOG04241	K	Transcription	RpoA/Rpc	DNA-directed RNA polymerase subunit D	COG00202	pfam01193	cd07030					1	1
arCOG02038	K	Transcription	-	Sugar-specific transcriptional regulator TrmB	COG01378						1		
arCOG02037	K	Transcription	-	Sugar-specific transcriptional regulator TrmB	COG01378	pfam01978						1	1
arCOG01057	K	Transcription	HxlR	DNA-binding transcriptional regulator, HxlR family	COG01733	pfam01638	cd00090						3
arCOG00608	K	Transcription	-	Predicted transcriptional regulator with C-terminal CBS domains	COG03620	pfam01381,pfa	cd00093,cd046	TIGR03070,TIGR01137					1
arCOG04248	K	Transcription	SIR2	NAD-dependent protein deacetylase, SIR2 family	COG00846	pfam02146	cd01413						1
arCOG04377	K	Transcription	-	Transcriptional regulator MarR family, contains HTH domain	COG04738								1
arCOG00826	K	Transcription	WecD	Acetyltransferase (GNAT) family	COG00454	pfam00583	cd04301	TIGR01575				1	
arCOG02644	K	Transcription	AcrR	Transcriptional regulator, TetR/AcrR family	COG01309	pfam00440		TIGR03613					1
arCOG05152	K	Transcription	-	Transcriptional regulator, contains HTH domain									1
arCOG02271	K	Transcription	-	Transcriptional regulator, contains HTH domain	COG03413	pfam04967						4	2
arCOG02274	K	Transcription	-	Transcriptional regulator, contains HTH domain	COG03413	pfam04967							3
arCOG02280	K	Transcription	-	Transcriptional regulator, contains HTH domain	COG03413	pfam04967						1	
arCOG04818	K	Transcription	HepA	Superfamily II DNA/RNA helicase, SNF2 family	COG00553	pfam13091,pfa	cd09178,cd000	TIGR01587				1	
arCOG02197	J	Translation, ribosomal structure and biogenesis	Cgi121	Subunit of KEOPS complex (Cgi121BUD32KAE1)	COG01617	pfam08617			5	1			
arCOG04107	J	Translation, ribosomal structure and biogenesis	SUI2	Translation initiation factor 2, alpha subunit (eIF-2alpha)	COG01093	pfam00575,pfa	cd04452	TIGR00717	5	1			
arCOG02466	J	Translation, ribosomal structure and biogenesis	GAR1	RNA-binding protein involved in rRNA processing	COG03277				4	1			
arCOG02286	J	Translation, ribosomal structure and biogenesis	Ncl1	Ribosome biogenesis protein, NOL1/NOP2/fmu family	COG03270	pfam13636			4	1			
arCOG04108	J	Translation, ribosomal structure and biogenesis	RPS27A	Ribosomal protein S27E	COG02051	pfam01667			4			1	
arCOG00973	J	Translation, ribosomal structure and biogenesis	RsmB	16S rRNA C967 or C1407 C5-methylase, RsmB/RsmF family	COG00144	pfam01189		TIGR00446	4	1		1	
arCOG00042	J	Translation, ribosomal structure and biogenesis	TilS	tRNA(Ile)-lysidine synthase TilS/MesJ	COG00037	pfam01171	cd01993	TIGR02432	4			1	
arCOG00033	J	Translation, ribosomal structure and biogenesis	Trm5	Wybutosine (yW) biosynthesis enzyme, Trm5 methyltransferase	COG02520	pfam02475			4		1	2	1
arCOG04449	J	Translation, ribosomal structure and biogenesis	TruA	Pseudouridylate synthase	COG00101	pfam01416,pfa	cd02866	TIGR00071	4			1	
arCOG01179	J	Translation, ribosomal structure and biogenesis	InfA	Translation initiation factor 1 (IF-1)	COG00361	pfam01176	cd05793	TIGR00523	3				
arCOG04150	J	Translation, ribosomal structure and biogenesis	Krr1	rRNA processing protein Krr1/Pno1, contains KH domain	COG01094			TIGR03665	3				
arCOG04149	J	Translation, ribosomal structure and biogenesis	NMD3	NMD protein affecting ribosome stability and mRNA decay	COG01499	pfam04981			3	1		1	1
arCOG00078	J	Translation, ribosomal structure and biogenesis	NOP1	Fibrillarin-like rRNA methylase	COG01889	pfam01269			3			1	1
arCOG00784	J	Translation, ribosomal structure and biogenesis	POP4	RNase P/RNase MRP subunit p29	COG01588	pfam01868			3		1	2	
arCOG00501	J	Translation, ribosomal structure and biogenesis	Rnz	Ribonuclease Z, beta-lactamase superfamily hydrolase	COG01234	pfam12706		TIGR02651	3			2	1
arCOG04070	J	Translation, ribosomal structure and biogenesis	RplC	Ribosomal protein L3	COG00087	pfam00297		TIGR03626	3				
arCOG00985	J	Translation, ribosomal structure and biogenesis	Tma20	Predicted RNA-binding protein, contains PUA domain	COG02016	pfam09183,pfam01472		TIGR03684	3	1	1	2	1
arCOG00047	J	Translation, ribosomal structure and biogenesis	Trm11	tRNA G10 N-methylase Trm11	COG01041	pfam01170		TIGR01177	3		1		
arCOG04156	J	Translation, ribosomal structure and biogenesis	TYW3	Wybutosine (yW) biosynthesis enzyme	COG01590	pfam02676			3				
arCOG00910	J	Translation, ribosomal structure and biogenesis	-	Predicted RNA methylase	COG02263	pfam13659			2	1	1	1	1
arCOG01695	J	Translation, ribosomal structure and biogenesis	-	Predicted component of the ribosome quality control (RQC) complex, YloA/Tl	COG01293	pfam05833,pfam05670			2	1	1	1	1
arCOG00676	J	Translation, ribosomal structure and biogenesis	Csl4	Exosome complex RNA-binding protein Csl4, contains S1 and Zn-ribbon doma	COG01096	pfam14382,pfa	cd05692		2	1	1	1	1
arCOG00809	J	Translation, ribosomal structure and biogenesis	LeuS	Leucyl-tRNA synthetase	COG00495	pfam00133,pfa	cd00812,cd008	TIGR00395	2				
arCOG00842	J	Translation, ribosomal structure and biogenesis	RimL	Acetyltransferase, RimL family	COG01670	pfam13302			2		1		
arCOG04345	J	Translation, ribosomal structure and biogenesis	RPR2	RNase P subunit RPR2	COG02023	pfam04032			2		1	1	
arCOG01701	J	Translation, ribosomal structure and biogenesis	SEN2	tRNA splicing endonuclease	COG01676	pfam02778,pfam01974		TIGR00324	2	1	1		
arCOG10124	J	Translation, ribosomal structure and biogenesis	TYW2	Wybutosine (yW) biosynthesis enzyme, TYW2 transferase	COG02520	pfam02475	cd02440	TIGR01444	2			2	1
arCOG01358	J	Translation, ribosomal structure and biogenesis	MiaB	2-methylthioadenine synthetase	COG00621	pfam00919,pfa	cd01335	TIGR00089	1				
arCOG00721	J	Translation, ribosomal structure and biogenesis	Nob1	Endonuclease Nob1, consists of a PIN domain and a Zn-ribbon module	COG01439		cd09876		1	1		1	1
arCOG00038	J	Translation, ribosomal structure and biogenesis	ThiI	tRNA S(4)U 4-thiouridine synthase	COG00301	pfam02926,pfa	cd11716,cd017	TIGR00342,TIGR04271	1		1	1	1
arCOG00990	J	Translation, ribosomal structure and biogenesis	-	Queuine tRNA-ribosyltransferase, contain PUA domain	COG01549	pfam14810,pfam01472		TIGR00432		1		2	1
arCOG01346	J	Translation, ribosomal structure and biogenesis	-	RNA-binding protein containing KH domain, possibly ribosomal protein	COG01534	pfam01985						2	1
arCOG04130	J	Translation, ribosomal structure and biogenesis	-	Predicted RNA-binding protein	COG01491	pfam04919						1	1
arCOG01254	J	Translation, ribosomal structure and biogenesis	AlaX	Ser-tRNA(Ala) deacylase AlaX (editing enzyme)	COG02872	pfam01411,pfam07973		TIGR03683				2	1
arCOG00487	J	Translation, ribosomal structure and biogenesis	ArgS	Arginyl-tRNA synthetase	COG00018	pfam00750,pfa	cd00671,cd079	TIGR00456				1	1
arCOG01185	J	Translation, ribosomal structure and biogenesis	Bud32	tRNA A-37 threonylcarbamoyl transferase component Bud32	COG03642	pfam00069	cd05151	TIGR03724				3	
arCOG04249	J	Translation, ribosomal structure and biogenesis	CCA1	tRNA nucleotidyltransferase (CCA-adding enzyme)	COG01746	pfam01909,pfa	cd05400	TIGR03671		1	1		
arCOG00486	J	Translation, ribosomal structure and biogenesis	CysS	Cysteinyl-tRNA synthetase	COG00215	pfam01406,pfa	cd00672,cd079	TIGR00435		1			
arCOG04161	J	Translation, ribosomal structure and biogenesis	DPH5	Diphthamide biosynthesis methyltransferase	COG01798	pfam00590	cd11647	TIGR00522					1
arCOG00035	J	Translation, ribosomal structure and biogenesis	Dph6	Diphthamide synthase (EF-2-diphthine--ammonia ligase)	COG02102	pfam01902	cd01994	TIGR03679				1	
arCOG04332	J	Translation, ribosomal structure and biogenesis	EbsC	Cys-tRNA(Pro)/Cys-tRNA(Cys) deacylase, ybaK family	COG02606	pfam04073	cd04333	TIGR00011				2	
arCOG01988	J	Translation, ribosomal structure and biogenesis	EFB1	Translation elongation factor EF-1beta	COG02092	pfam00736	cd00292	TIGR00489					1
arCOG04277	J	Translation, ribosomal structure and biogenesis	Efp	Translation elongation factor P (EF-P)/translation initiation factor 5A (eIF-5A)	COG00231	pfam08207,pfa	cd04467	TIGR00037					1
arCOG04312	J	Translation, ribosomal structure and biogenesis	Fcf1	rRNA-processing protein FCF1, contains PIN domain	COG01412		cd09879					2	1
arCOG01559	J	Translation, ribosomal structure and biogenesis	FusA	Translation elongation factor G, EF-G (GTPase)	COG00480	pfam00009,pfa	cd01885,cd037	TIGR00490				1	
arCOG00978	J	Translation, ribosomal structure and biogenesis	GCD14	tRNA(1-methyladenosine) methyltransferase	COG02519	pfam08704	cd02440	TIGR03534		1	1	1	1
arCOG01616	J	Translation, ribosomal structure and biogenesis	GEK1	D-aminoacyl-tRNA deacylase, involved in ethanol tolerance	COG01650	pfam04414							1
arCOG00109	J	Translation, ribosomal structure and biogenesis	HemK	Methylase of polypeptide chain release factors	COG02890	pfam13659	cd02440	TIGR00537			1		
arCOG00404	J	Translation, ribosomal structure and biogenesis	HisS	Histidyl-tRNA synthetase	COG00124	pfam13393,pfa	cd00773,cd008	TIGR00442		1			1
arCOG00807	J	Translation, ribosomal structure and biogenesis	IleS	Isoleucyl-tRNA synthetase	COG00060	pfam00133,pfa	cd00818,cd008	TIGR00392				1	
arCOG01560	J	Translation, ribosomal structure and biogenesis	InfB	Translation initiation factor 2 (IF-2; GTPase)	COG00532	pfam00009,pfa	cd01887,cd037	TIGR00491					1
arCOG01736	J	Translation, ribosomal structure and biogenesis	LigT	2'-5' RNA ligase	COG01514	pfam13563		TIGR02258			1		
arCOG00810	J	Translation, ribosomal structure and biogenesis	MetG	Methionyl-tRNA synthetase	COG00143	pfam09334,pfa	cd00814,cd079	TIGR00398				1	
arCOG00906	J	Translation, ribosomal structure and biogenesis	Nop10	rRNA maturation protein Nop10, contains Zn-ribbon domain	COG02260	pfam04135						1	

arCOG01741	J	Translation, ribosomal structure and biogenesis	PelA	Release factor eRF1	COG01537	pfam03463,pfam03464,pfam03		TIGR00111		1		2	1
arCOG00412	J	Translation, ribosomal structure and biogenesis	PheT	Phenylalanyl-tRNA synthetase beta subunit	COG00072	pfam03483,pfa	cd00769	TIGR00471				2	1
arCOG00402	J	Translation, ribosomal structure and biogenesis	ProS	Prolyl-tRNA synthetase	COG00442	pfam00587,pfa	cd00778,cd008	TIGR00408				1	
arCOG04228	J	Translation, ribosomal structure and biogenesis	Pth2	Peptidyl-tRNA hydrolase	COG01990	pfam01981	cd02430	TIGR00283				2	2
arCOG01015	J	Translation, ribosomal structure and biogenesis	Pus10	tRNA U54 and U55 pseudouridine synthase Pus10	COG01258			TIGR01213				1	1
arCOG00039	J	Translation, ribosomal structure and biogenesis	QueC	7-cyano-7-deazaguanine synthase (queuosine biosynthesis)	COG00603	pfam06508	cd01995	TIGR00364			1		
arCOG05111	J	Translation, ribosomal structure and biogenesis	RlmH	23S rRNA pseudoU1915 N3-methylase RlmH	COG01576	pfam02590		TIGR00246					1
arCOG06660	J	Translation, ribosomal structure and biogenesis	RluA	Pseudouridylate synthase, 23S RNA-specific	COG00564	pfam00849	cd02869	TIGR00005				1	1
arCOG00780	J	Translation, ribosomal structure and biogenesis	RPL18A	Ribosomal protein L18E	COG01727	pfam00828		TIGR01071			1		1
arCOG04175	J	Translation, ribosomal structure and biogenesis	RPL20A	Ribosomal protein L20A (L18A)	COG02157	pfam01775							1
arCOG01950	J	Translation, ribosomal structure and biogenesis	RPL24A	Ribosomal protein L24E	COG02075	pfam01246	cd00472				1		
arCOG01752	J	Translation, ribosomal structure and biogenesis	RPL30	Ribosomal protein L30E	COG01911	pfam01248				1			
arCOG04049	J	Translation, ribosomal structure and biogenesis	RPL40A	Ribosomal protein L40E	COG01552							1	
arCOG04208	J	Translation, ribosomal structure and biogenesis	RPL43A	Ribosomal protein L37AE/L43A	COG01997	pfam01780		TIGR00280				1	
arCOG04288	J	Translation, ribosomal structure and biogenesis	RplJ	Ribosomal protein L10	COG00244	pfam00466	cd05795					1	
arCOG04372	J	Translation, ribosomal structure and biogenesis	RplK	Ribosomal protein L11	COG00080	pfam03946,pfa	cd00349	TIGR01632				2	1
arCOG04094	J	Translation, ribosomal structure and biogenesis	RplX	Ribosomal protein L24	COG00198		cd06089	TIGR01080				2	
arCOG04186	J	Translation, ribosomal structure and biogenesis	RPS1A	Ribosomal protein S3AE	COG01890	pfam01015				1		1	
arCOG04182	J	Translation, ribosomal structure and biogenesis	RPS24A	Ribosomal protein S24E	COG02004	pfam01282				1			1
arCOG04239	J	Translation, ribosomal structure and biogenesis	RpsD	Ribosomal protein S4 or related protein	COG00522	pfam00163,pfa	cd00165	TIGR01018				1	1
arCOG00678	J	Translation, ribosomal structure and biogenesis	Rrp4	Exosome complex RNA-binding protein Rrp4, contains S1 and KH domains	COG01097		cd05789			1			1
arCOG04131	J	Translation, ribosomal structure and biogenesis	RsmA	16S rRNA A1518 and A1519 N6-dimethyltransferase RsmA/KsgA/DIM1	COG00030	pfam00398		TIGR00755				1	1
arCOG04246	J	Translation, ribosomal structure and biogenesis	RtcB	RNA 3'-P ligase, RtcB family protein	COG01690	pfam01139		TIGR03073		1			
arCOG01564	J	Translation, ribosomal structure and biogenesis	SelB	Selenocysteine-specific translation elongation factor or SelB-II domain	COG03276		cd03696	TIGR00475		1			
arCOG01923	J	Translation, ribosomal structure and biogenesis	SIK1	RNA processing factor Prp31, contains Nop domain	COG01498	pfam01798					1	1	1
arCOG01952	J	Translation, ribosomal structure and biogenesis	SUA5	tRNA A37 threonylcarbamoyladenosine synthetase subunit TsaC/SUA5/YrdC	COG00009	pfam01300		TIGR00057		1		1	1
arCOG04223	J	Translation, ribosomal structure and biogenesis	SUI1	Translation initiation factor 1 (eIF-1/SUI1)	COG00023	pfam01253	cd11567	TIGR01158				1	
arCOG00056	J	Translation, ribosomal structure and biogenesis	Tan1	tRNA(Ser,Leu) C12 N-acetylase TAN1, contains THUMP domain	COG01818	pfam02926,pfa	cd11718,cd061	TIGR00342		1			
arCOG01630	J	Translation, ribosomal structure and biogenesis	TdcF	Translation initiation inhibitor, yjgF family	COG00251	pfam01042	cd00448	TIGR00004			1	2	1
arCOG01561	J	Translation, ribosomal structure and biogenesis	TEF1	Translation elongation factor EF-1 alpha, GTPase	COG05256	pfam00009,pfa	cd01883,cd036	TIGR00483				3	
arCOG00401	J	Translation, ribosomal structure and biogenesis	ThrS	Threonyl-tRNA synthetase	COG00441	pfam00587,pfa	cd00771,cd008	TIGR00418		1			
arCOG01115	J	Translation, ribosomal structure and biogenesis	TiaS	tRNA(Ile2) 2-agmatinylcytidine synthetase; containing Zn-ribbon domain and	COG01571	pfam08489,pfa	cd04482	TIGR03280					1
arCOG04176	J	Translation, ribosomal structure and biogenesis	TIF6	Translation initiation factor 6 (eIF-6)	COG01976	pfam01912	cd00527	TIGR00323		1			
arCOG01219	J	Translation, ribosomal structure and biogenesis	TRM1	N2,N2-dimethylguanosine tRNA methyltransferase	COG01867	pfam02005		TIGR00308				1	1
arCOG01018	J	Translation, ribosomal structure and biogenesis	TrmJ	tRNA C32,U32 (ribose-2'-O)-methylase TrmJ or a related methyltransferase	COG00565	pfam00588		TIGR00050				1	2
arCOG00987	J	Translation, ribosomal structure and biogenesis	TruB	Pseudouridine synthase	COG00130	pfam08068,pfa	cd02572	TIGR00425		1		1	1
arCOG04252	J	Translation, ribosomal structure and biogenesis	TruD	tRNA(Glu) U13 pseudouridine synthase TruD	COG00585	pfam01142	cd02577	TIGR00094		1		1	1
arCOG04733	J	Translation, ribosomal structure and biogenesis	Tsr3	Ribosome biogenesis protein Tsr3, contains Fer4-like metal-binding domain c	COG02042	pfam04034						1	
arCOG00541	J	Translation, ribosomal structure and biogenesis	YSH1	Predicted exonuclease of the beta-lactamase fold involved in RNA processing	COG01236	pfam00753,pfam10996,pfam01		TIGR03675				4	
CELLULAR PROCESSES AND SIGNALING													
arCOG11012	D	Cell cycle control, cell division, chromosome partitioning	ATS1	Alpha-tubulin suppressor and related RCC1 domain-containing proteins	COG05184	pfam13540,pfam00415,pfam00415,pfam00415,pfam00415,pfam00415						1	2
arCOG02201	D	Cell cycle control, cell division, chromosome partitioning	FtsZ	Cell division GTPase	COG00206	pfam00091,pfa	cd02201	TIGR00065	5				1
arCOG05007	D	Cell cycle control, cell division, chromosome partitioning	Maf	Nucleotide-binding protein implicated in inhibition of septum formation	COG00424	pfam02545	cd00555	TIGR00172				1	1
arCOG03061	D	Cell cycle control, cell division, chromosome partitioning	MreB	Actin-like ATPase involved in cell morphogenesis	COG01077	pfam06723	cd10227	TIGR00904					3
arCOG00586	D	Cell cycle control, cell division, chromosome partitioning	Soj	ATPase involved in chromosome partitioning, ParA family	COG01192	pfam01656	cd02042	TIGR03453				2	
arCOG02416	N	Cell motility	-	Pilin/Flagellin, FlaG/FlaF family	COG03430	pfam07790							2
arCOG02079	N	Cell motility	-	S-layer protein, possibly associated with type IV pili like system	COG01361				4				
arCOG01829	N	Cell motility	FlaB	Archaeal flagellins	COG01681							1	
arCOG01822	N	Cell motility	FlaG	Archaellum protein G, flagellin of FlaG/FlaF family	COG03354								1
arCOG04148	N	Cell motility	FlaH	ATPase involved in biogenesis of archaellum	COG02874	pfam06745	cd01394	TIGR03881					1
arCOG01809	N	Cell motility	FlaJ	Archaellum assembly protein J, TadC family	COG01955								1
arCOG02298	N	Cell motility	FlaK/PulO	Peptidase A24A, prepilin type IV	COG01989	pfam01478,pfam06847				1			1
arCOG01817	N	Cell motility	VirB11	ATPase involved in archaellum/pili biosynthesis	COG00630	pfam00437	cd01130	TIGR03819					1
arCOG01410	M	Cell wall/membrane/envelope biogenesis	AglA	Glycosyltransferase	COG00438	pfam13579,pfa	cd03794	TIGR02149	9	2			1
arCOG00118	M	Cell wall/membrane/envelope biogenesis	WecE	Predicted pyridoxal phosphate-dependent enzyme apparently involved in reg	COG00399	pfam01041	cd00616	TIGR03588	7				1
arCOG00894	M	Cell wall/membrane/envelope biogenesis	-	Glycosyl transferase family 2	COG00463	pfam00535	cd04179	TIGR04182	5			1	
arCOG00666	M	Cell wall/membrane/envelope biogenesis	GCD1	N-acetylglucosamine-1-phosphate uridyltransferase	COG01208	pfam00483,pfa	cd04181,cd056	TIGR03992	4		1	1	1
arCOG01568	M	Cell wall/membrane/envelope biogenesis	MscS	Small-conductance mechanosensitive channel	COG00668	pfam00924			4			5	2
arCOG01411	M	Cell wall/membrane/envelope biogenesis	RfaG	Glycosyltransferase	COG00438	pfam13439,pfa	cd03801	TIGR03999	4			3	1
arCOG00252	M	Cell wall/membrane/envelope biogenesis	WecC	UDP-N-acetyl-D-mannosaminuronate dehydrogenase	COG00677	pfam03721,pfa	cd05266	TIGR03026	4				
arCOG03336	M	Cell wall/membrane/envelope biogenesis	-	Surface protein containing fasciclin-like repeats	COG02335	pfam02469			3				1
arCOG01403	M	Cell wall/membrane/envelope biogenesis	AglL	Glycosyltransferase	COG00438	pfam13439,pfa	cd03804		3			5	2
arCOG02209	M	Cell wall/membrane/envelope biogenesis	AglR	MATE family membrane protein, Rfbx family	COG02244		cd13128		3	1			
arCOG00899	M	Cell wall/membrane/envelope biogenesis	AglD2	Predicted flippase	COG00392	pfam03706		TIGR00374	2				
arCOG00668	M	Cell wall/membrane/envelope biogenesis	GCD1	Nucleoside-diphosphate-sugar pyrophosphorylase involved in lipopolysaccha	COG01208	pfam00483,pfa	cd04181,cd056	TIGR03992	2				
arCOG06759	M	Cell wall/membrane/envelope biogenesis	RfaG	Glycosyltransferase	COG00438				2				
arCOG01222	M	Cell wall/membrane/envelope biogenesis	TagD	Cytidylyltransferase fused to conserved domain of DUF357 family	COG00615	pfam01467	cd02170	TIGR02199	2				
arCOG03335	M	Cell wall/membrane/envelope biogenesis	-	Surface protein containing fasciclin-like repeats	COG02335	pfam02469			1				
arCOG03199	M	Cell wall/membrane/envelope biogenesis	AglH/Rfe	UDP-N-acetylmuramyl pentapeptide phosphotransferase/UDP-N-acetylgluco	COG00472	pfam00953	cd06856	TIGR00445	1				
arCOG02820	M	Cell wall/membrane/envelope biogenesis	MurE	UDP-N-acetylmuramyl tripeptide synthase	COG00769	pfam08245,pfam02875		TIGR01082	1				
arCOG04827	M	Cell wall/membrane/envelope biogenesis	TagB	Glycosyl/glycerophosphate transferase involved in teichoic acid biosynthesis	COG01887	pfam04464			1				
arCOG01369	M	Cell wall/membrane/envelope biogenesis	WcaG	Nucleoside-diphosphate-sugar epimerase	COG00451	pfam01370	cd05234	TIGR01179	1	1		2	4
arCOG01392	M	Cell wall/membrane/envelope biogenesis	WecB	UDP-N-acetylglucosamine 2-epimerase	COG00381	pfam02350	cd03786	TIGR00236	1				
arCOG02488	M	Cell wall/membrane/envelope biogenesis	-	Cell surface protein, a component of a putative secretion system	SC.00293					1			
arCOG01385	M	Cell wall/membrane/envelope biogenesis	-	Glycosyl transferase family 2	COG00463	pfam00535	cd02522	TIGR04283			1	1	2
arCOG02532	M	Cell wall/membrane/envelope biogenesis	-	Secreted Beta-propeller repeat protein fused to CARDB-like adhesion module	COG03291	pfam07705		TIGR04213		1	2		2
arCOG07560	M	Cell wall/membrane/envelope biogenesis	-	Cell surface protein							1	1	
arCOG00895	M	Cell wall/membrane/envelope biogenesis	-	Glycosyl transferase family 2	COG00463	pfam00535	cd06442	TIGR04182					1
arCOG02487	M	Cell wall/membrane/envelope biogenesis	-	Cell surface protein, a component of a putative secretion system		pfam02369							1
arCOG02538	M	Cell wall/membrane/envelope biogenesis	-	Secreted protein, with PKD repeat domain	COG03291			TIGR04213					1
arCOG00896	M	Cell wall/membrane/envelope biogenesis	-	Glycosyl transferase family 2	COG00463	pfam00535	cd04179	TIGR04182					1
arCOG05365	M	Cell wall/membrane/envelope biogenesis	AglB	Oligosaccharyltransferase membrane subunit	COG01287	pfam02516,pfam13620,pfam13		TIGR04154		1	1	1	2
arCOG00664	M	Cell wall/membrane/envelope biogenesis	AglF/RfbA	Glucose-1-phosphate uridylyltransferase	COG01209	pfam00483	cd04181	TIGR03992					1
arCOG00253	M	Cell wall/membrane/envelope biogenesis	AglM/Ugd	UDP-glucose 6-dehydrogenase	COG01004	pfam03721,pfam00984,pfam03		TIGR03026					1
arCOG01381	M	Cell wall/membrane/envelope biogenesis	AglO	Glycosyl transferase family 2	COG00463	pfam00535	cd00761	TIGR04283				1	1
arCOG00663	M	Cell wall/membrane/envelope biogenesis	GCD1	Nucleoside-diphosphate-sugar pyrophosphorylase involved in lipopolysaccha	COG01208	pfam00483	cd04181	TIGR03992		1			
arCOG00057	M	Cell wall/membrane/envelope biogenesis	GlmS	Glucosamine 6-phosphate synthetase	COG00449	pfam00310,pfa	cd00714,cd050	TIGR01135		1		1	1
arCOG01373	M	Cell wall/membrane/envelope biogenesis	Gmd	GDP-D-mannose dehydratase	COG01089	pfam01370,pfa	cd05260	TIGR01472					1
arCOG02313	M	Cell wall/membrane/envelope biogenesis	LolE	ABC-type transport system, involved in lipoprotein release, permease compo	COG04591	pfam12704,pfam02687		TIGR02212					1
arCOG02821	M	Cell wall/membrane/envelope biogenesis	MurD	UDP-N-acetylmuramoylalanine-D-glutamate ligase	COG00771	pfam08245,pfam02875		TIGR01082		1			
arCOG07536	M	Cell wall/membrane/envelope biogenesis	NanM	N-acetylneuraminic acid mutarotase	COG03055								1
arCOG01405	M	Cell wall/membrane/envelope biogenesis	RfaG	Glycosyltransferase	COG00438	pfam13439,pfa	cd03794	TIGR04063				1	
arCOG01409	M	Cell wall/membrane/envelope biogenesis	RfaG	Glycosyltransferase	COG00438	pfam13692	cd03794	TIGR03087				1	
arCOG00194	V	Defense mechanisms	CcmA	ABC-type multidrug transport system, ATPase component	COG01131	pfam00005	cd03230	TIGR01188	9	2			1
arCOG00786	V	Defense mechanisms	Cas4	CRISPR-associated protein Cas4, RecB family exonuclease	COG01468	pfam01930	cd09637	TIGR00372	5	1			1
arCOG01731	V	Defense mechanisms	NorM	Na+-driven multidrug efflux pump	COG00534	pfam01554,pfa	cd13137	TIGR00797	5				
arCOG01463	V	Defense mechanisms	-	ABC-type multidrug transport system, permease component	COG00842	pfam01061		TIGR01247	4				
arCOG02312	V	Defense mechanisms	SalY	ABC-type antimicrobial peptide transport system, permease component	COG00577	pfam02687		TIGR02212	4			1	1
arCOG03787	V	Defense mechanisms	-	LEA14-like dessication related protein	COG05608	pfam03168,pfam03168			3	1			
arCOG00922	V	Defense mechanisms	SalX	ABC-type antimicrobial peptide transport system, ATPase component	COG01136	pfam00005	cd03255	TIGR02673	3				1
arCOG00196	V	Defense mechanisms	CcmA	ABC-type multidrug transport system, ATPase component	COG01131	pfam00005	cd03230	TIGR03740	2		1		
arCOG03779	V	Defense mechanisms	McrB	GTPase subunit of restriction endonuclease	COG01401	pfam07728	cd00009		2				
arCOG00516	V	Defense mechanisms	NimA	Nitroimidazol reductase NimA or a related FMN-containing flavoprotein, pyri	COG03467	pfam01243		TIGR04023	2				
arCOG00525	V	Defense mechanisms	NimA	Nitroimidazol reductase NimA or a related FMN-containing flavoprotein, pyri	COG03467	pfam12724,pfam01243			2				
arCOG03966	V	Defense mechanisms	-	CopG family DNA-binding protein							1		
arCOG03166	V	Defense mechanisms	-	AAA+ superfamily ATPase fused to HTH and RecB nuclease domains	COG01672							1	1
arCOG03521	V	Defense mechanisms	-	Type II restriction enzyme, methylase subunit	COG01002	pfam12950							1
arCOG02632	V	Defense mechanisms	HsdM	Type I restriction-modification system methyltransferase subunit	COG00286	pfam12161,pfam02384		TIGR00497				2	
arCOG02777	V	Defense mechanisms	Mrr	Restriction endonuclease	COG01715	pfam04471							1
arCOG01663	V	Defense mechanisms	RelE	Cytotoxic translational repressor of toxin-antitoxin stability system	COG02026	pfam05016					1		
arCOG02204	U	Intracellular trafficking, secretion, and vesicular transport	Sss1	Preprotein translocase subunit Sss1	COG02443				4	1			
arCOG01919	U	Intracellular trafficking, secretion, and vesicular transport	TatC	Sec-independent protein secretion pathway component TatC	COG00805	pfam00902		TIGR00945	3	1		3	1
arCOG01217	U	Intracellular trafficking, secretion, and vesicular transport	SEC65	Signal recognition particle 19 kDa protein	COG01400				2	1	1	2	1
arCOG02694	U	Intracellular trafficking, secretion, and vesicular transport	TatA	Sec-independent protein secretion pathway component	COG01826			TIGR01411	2			1	
arCOG02673	U	Intracellular trafficking, secretion, and vesicular transport	-	OxaA/SpoJ/YidC translocase/secretase, sec-independent itegration of nascen	COG01422	pfam01956						1	1
arCOG04471	U	Intracellular trafficking, secretion, and vesicular transport	-	Exosortase	COG04083	pfam09721		TIGR04125					1
arCOG01227	U	Intracellular trafficking, secretion, and vesicular transport	FtsY	Signal recognition particle GTPase	COG00552	pfam02881,pfa	cd03115	TIGR00064				4	1
arCOG03383	U	Intracellular trafficking, secretion, and vesicular transport	TolB	Periplasmic component of the Tol biopolymer transport system	COG00823	pfam00930,pfam08308		TIGR02800					1
arCOG01241	X	Mobilome: prophages, transposons	XerC	XerD/XerC family integrase	COG00582	pfam13495,pfa	cd00796	TIGR02224				2	1
arCOG03473	X	Mobilome: prophages, transposons	-	Transposase	COG05421				2				
arCOG03669	O	Posttranslational modification, protein turnover, chaperones	-	Subtilase family protease	COG04934	pfam09286	cd11377,cd04056		12	2	3	3	
arCOG01912	O	Posttranslational modification, protein turnover, chaperones	-	Membrane protein implicated in regulation of membrane protease activity	COG01585	pfam01957			3				
arCOG02815	O	Posttranslational modification, protein turnover, chaperones	-	Conserved domain frequently associated with peptide methionine sulfoxide r	COG00229	pfam01641		TIGR00357	3				
arCOG03607	O	Posttranslational modification, protein turnover, chaperones	-	Cysteine protease, C1A family	COG04870	pfam00112	cd02619		2	1			
arCOG02007	O	Posttranslational modification, protein turnover, chaperones	-	Prenyltransferase family protein containing thioredoxin domain	COG01331	pfam03190	cd02955		2		1		
arCOG03614	O	Posttranslational modification, protein turnover, chaperones	-	Cysteine protease, C1A family	COG04870	pfam00112	cd02619			1			
arCOG03202	O	Posttranslational modification, protein turnover, chaperones	-	Collagenase family protease	COG00826	pfam01297,pfam01136					1		
arCOG10598	O	Posttranslational modification, protein turnover, chaperones	-	ATP-dependent protease La PUA-like domain		pfam02190						1	1
arCOG03685	O	Posttranslational modification, protein turnover, chaperones	-	Predicted redox protein, regulator of disulfide bond formation	COG01765	pfam02566							1
arCOG00312	O	Posttranslational modification, protein turnover, chaperones	AhpC	AhpC/TSA family peroxiredoxin	COG00450	pfam00578,pfa	cd03016	TIGR03137	5				1
arCOG01297	O	Posttranslational modification, protein turnover, chaperones	AhpF	Alkyl hydroperoxide reductase, large subunit	COG03634	pfam07992		TIGR03140	1				
arCOG00702	O	Posttranslational modification, protein turnover, chaperones	AprE	Subtilisin-like serine protease	COG01404	pfam00082	cd07477	TIGR03921	6	1	1	2	1
arCOG02553	O	Posttranslational modification, protein turnover, chaperones	AprE	Subtilisin-like serine protease	COG01404	pfam05048			2	1	1		
arCOG06823	O	Posttranslational modification, protein turnover, chaperones	AprE	Subtilisin-like serine protease	COG01404	pfam00082	cd05562	TIGR03921	1			1	
arCOG03610	O	Posttranslational modification, protein turnover, chaperones	AprE	Subtilisin-like serine protease	COG01404		cd02619				2	1	
arCOG01226	O	Posttranslational modification, protein turnover, chaperones	ArgK	Putative periplasmic protein kinase ArgK or related GTPase of G3E family	COG01703	pfam03308	cd03114	TIGR00750		1		2	1
arCOG00310	O	Posttranslational modification, protein turnover, chaperones	Bcp	Peroxiredoxin	COG01225	pfam00578	cd03018	TIGR03137	2			7	
arCOG01328	O	Posttranslational modification, protein turnover, chaperones	CcmB	ABC-type transport system involved in cytochrome c biogenesis, permease co	COG02386				2				
arCOG00267	O	Posttranslational modification, protein turnover, chaperones	CcmC	ABC-type transport system involved in cytochrome c biogenesis, permease co	COG00755	pfam01578		TIGR01191		1			
arCOG04957	O	Posttranslational modification, protein turnover, chaperones	CcmE	Cytochrome c-type biogenesis protein CcmE	COG02332					1		2	1
arCOG00270	O	Posttranslational modification, protein turnover, chaperones	CcmF	Cytochrome c biogenesis factor	COG01138	pfam01578		TIGR00353				1	1

arCOG01308	O	Posttranslational modification, protein turnover, chaperones	Cdc48	ATPase of the AAA+ class , CDC48 family	COG0464	pfam00004,pfa	cd00009,cd000	TIGR01243				3	1
arCOG00065	O	Posttranslational modification, protein turnover, chaperones	csdA	Selenocysteine lyase/Cysteine desulfurase	COG0520	pfam00266	cd06453	TIGR01979					1
arCOG03103	O	Posttranslational modification, protein turnover, chaperones	CtaA	Uncharacterized protein required for cytochrome oxidase assembly	COG01612	pfam02628				1			1
arCOG06880	O	Posttranslational modification, protein turnover, chaperones	DnaJ	DnaJ type Zn finger domain	COG00484			TIGR02349				3	
arCOG02846	O	Posttranslational modification, protein turnover, chaperones	DnaJ	DnaJ-class molecular chaperone	COG00484	pfam00226	cd06257	TIGR02349					1
arCOG04142	O	Posttranslational modification, protein turnover, chaperones	DYS1	Deoxyhypusine synthase	COG01899	pfam01916		TIGR00321		1			
arCOG04712	O	Posttranslational modification, protein turnover, chaperones	ECM4	Predicted glutathione S-transferase	COG00435	pfam13409,pfa	cd03190						1
arCOG01341	O	Posttranslational modification, protein turnover, chaperones	GIM5	Predicted prefoldin, molecular chaperone implicated in de novo protein folding	COG1730	pfam02996	cd00584	TIGR00293		1	1		
arCOG05154	O	Posttranslational modification, protein turnover, chaperones	GroEL	Chaperonin GroEL (HSP60 family)	COG00459	pfam00118	cd03344	TIGR02348	5				
arCOG01257	O	Posttranslational modification, protein turnover, chaperones	GroEL	Chaperonin GroEL, HSP60 family	COG00459	pfam00118	cd03343	TIGR02339				7	2
arCOG04772	O	Posttranslational modification, protein turnover, chaperones	GrpE	Molecular chaperone GrpE (heat shock protein)	COG00576	pfam01025	cd00446			1		2	1
arCOG02606	O	Posttranslational modification, protein turnover, chaperones	GrxC	Glutaredoxin	COG00695	pfam00462	cd02976	TIGR02196	1				
arCOG02607	O	Posttranslational modification, protein turnover, chaperones	GrxC	Glutaredoxin	COG00695	pfam00462	cd02976	TIGR02196		1			1
arCOG01915	O	Posttranslational modification, protein turnover, chaperones	HflC	Membrane protease subunit, stomatin/prohibitin homolog	COG00330	pfam01145	cd08826	TIGR01933				1	
arCOG02959	O	Posttranslational modification, protein turnover, chaperones	Iap	Zn-dependent amino- or carboxypeptidase, M28 family	COG02234	pfam04389	cd05643						3
arCOG04560	O	Posttranslational modification, protein turnover, chaperones	IscA	Fe-S cluster assembly iron-binding protein IscA	COG00316	pfam01521		TIGR00049				2	
arCOG02443	O	Posttranslational modification, protein turnover, chaperones	NosY	ABC-type transport system involved in multi-copper enzyme maturation, per	COG01277	pfam12679			1				
arCOG02438	O	Posttranslational modification, protein turnover, chaperones	NosY	ABC-type transport system involved in multi-copper enzyme maturation, per	COG01277	pfam12679				1			
arCOG01846	O	Posttranslational modification, protein turnover, chaperones	PaaD	Metal-sulfur cluster biosynthetic enzyme	COG02151	pfam01883,pfa	cd02037	TIGR02945,TIGR01969	2				
arCOG00976	O	Posttranslational modification, protein turnover, chaperones	Pcm	Protein-L-isoaspartate carboxylmethyltransferase	COG02518	pfam01135	cd02440	TIGR00080			1	1	
arCOG05850	O	Posttranslational modification, protein turnover, chaperones	Pcp	Pyrrolidone-carboxylate peptidase (N-terminal pyroglutamyl peptidase)	COG02039	pfam01470	cd00501	TIGR00504					1
arCOG04767	O	Posttranslational modification, protein turnover, chaperones	PpiB	Peptidyl-prolyl cis-trans isomerase (rotamase) - cyclophilin family	COG00652							1	1
arCOG01306	O	Posttranslational modification, protein turnover, chaperones	RPT1	ATP-dependent 26S proteasome regulatory subunit	COG01222	pfam00004	cd00009	TIGR01242				1	
arCOG00609	O	Posttranslational modification, protein turnover, chaperones	RseP	Membrane-associated protease RseP, regulator of RpoE activity in bacteria	COG00750	pfam02163	cd06160			1			
arCOG04064	O	Posttranslational modification, protein turnover, chaperones	RseP	Membrane-associated protease RseP, regulator of RpoE activity in bacteria	COG00750	pfam02163	cd06159	TIGR00054				1	
arCOG00981	O	Posttranslational modification, protein turnover, chaperones	SlpA	FKBP-type peptidyl-prolyl cis-trans isomerase 2	COG01047	pfam00254		TIGR00115	2				
arCOG00980	O	Posttranslational modification, protein turnover, chaperones	SlpA	FKBP-type peptidyl-prolyl cis-trans isomerase 2	COG01047	pfam00254				1	1		
arCOG02784	O	Posttranslational modification, protein turnover, chaperones	SRAP	Putative SOS response-associated peptidase YedK	COG02135	pfam02586							1
arCOG03580	O	Posttranslational modification, protein turnover, chaperones	STE14	Putative protein-S-isoprenylcysteine methyltransferase	COG02020	pfam04191			2				
arCOG01715	O	Posttranslational modification, protein turnover, chaperones	SufB	Cysteine desulfurase activator SufB	COG00719	pfam01458		TIGR01981	3	1		1	1
arCOG07441	O	Posttranslational modification, protein turnover, chaperones	SurA	Parvulin-like peptidyl-prolyl isomerase	COG00760	pfam00639		TIGR02933				1	1
arCOG00321	O	Posttranslational modification, protein turnover, chaperones	TldD	Zn-dependent protease, component of TldD/TldE system	COG00312	pfam01523							1
arCOG00322	O	Posttranslational modification, protein turnover, chaperones	TldE	Inactivated Zn-dependent protease, component of TldD/TldE system	COG00312	pfam01523						1	1
arCOG06181	O	Posttranslational modification, protein turnover, chaperones	TrxA	Thiol-disulfide isomerase or thioredoxin	COG00526	pfam00578	cd02966		3	1	1		
arCOG01972	O	Posttranslational modification, protein turnover, chaperones	TrxA	Thiol-disulfide isomerase or thioredoxin	COG00526	pfam00085	cd02947	TIGR01068	1				
arCOG01929	O	Posttranslational modification, protein turnover, chaperones	XdhC	Xanthine and CO dehydrogenase maturation factor, XdhC/CoxF family	COG01975	pfam13478						1	
arCOG02985	T	Signal transduction mechanisms	-	Membrane proteinase of CAAX superfamily, regulator of anti-sigma factor	COG02339	pfam13367			3	1	1	1	1
arCOG02967	T	Signal transduction mechanisms	-	NACHT family NTPase fused to HEAT repeats domain	COG5635	pfam13646,pfam13646			3		1		1
arCOG01180	T	Signal transduction mechanisms	RIO1	Serine/threonine protein kinase involved in cell cycle control	COG01718	pfam01163	cd05145	TIGR03724	3				
arCOG04820	T	Signal transduction mechanisms	-	SOUL heme-binding protein		pfam04832			3			1	
arCOG06193	T	Signal transduction mechanisms	-	Signal transduction histidine kinase and PAS domains	COG00642	pfam13426,pfa	cd00130,cd001	TIGR00229,TIGR00229,TIGR01386	2				
arCOG04425	T	Signal transduction mechanisms	Wzb	Protein-tyrosine-phosphatase	COG00394	pfam01451	cd00115	TIGR02689	1				
arCOG05374	T	Signal transduction mechanisms	-	GAF domain-containing protein	COG01956	pfam13185			1		1		
arCOG01171	T	Signal transduction mechanisms	RAD55	RecA-superfamily ATPase implicated in signal transduction	COG00467	pfam06745	cd01124	TIGR03877					2
arCOG04453	T	Signal transduction mechanisms	DisA_N	Diadenylate cyclase (c-di-AMP synthetase), DisA_N domain	COG01624	pfam02457							1
arCOG03799	T	Signal transduction mechanisms	AtoS	GAF, PAS/PAC domains containing signal transduction protein	COG02202	pfam05763						1	2
arCOG01143	T	Signal transduction mechanisms	ApaH	Serine/threonine protein phosphatase PP2A family	COG00639	pfam00149	cd00144						2
arCOG02391	T	Signal transduction mechanisms	CheY	Rec domain	COG00784	pfam00072	cd00156	TIGR02154				3	1
arCOG03413	T	Signal transduction mechanisms	CDC14	Protein-tyrosine phosphatase	COG02453	pfam00782	cd00047						1
arCOG03517	T	Signal transduction mechanisms	-	Formylglycine-generating sulfatase enzyme	COG01262	pfam12867,pfam03781		TIGR03440					1
arCOG01992	T	Signal transduction mechanisms	SixA	Phosphohistidine phosphatase SixA	COG02062	pfam00300	cd07067	TIGR00249				1	
arCOG06801	T	Signal transduction mechanisms	SrkA	Ser/Thr protein kinase RdoA involved in Cpx stress response, MazF antagonis	COG02334	pfam01636	cd05153					3	
METABOLISM													
arCOG02767	E	Amino acid transport and metabolism	-	Metal-dependent membrane protease, CAAX family	COG01266	pfam02517			3				
arCOG03600	E	Amino acid transport and metabolism	-	Transglutaminase-like cysteine protease	COG01305				2				
arCOG14875	E	Amino acid transport and metabolism	-	Aspartyl protease		pfam05618			2			1	
arCOG03611	E	Amino acid transport and metabolism	-	Peptidase C1A subfamily			cd08546,cd14254		2				2
arCOG07760	E	Amino acid transport and metabolism	-	Peptidase family C25		pfam01364	cd02258		1				
arCOG04333	E	Amino acid transport and metabolism	-	Aspartate/tyrosine/aromatic aminotransferase	COG00436	pfam00155	cd00609	TIGR03947		1			
arCOG01130	E	Amino acid transport and metabolism	-	Aspartate/tyrosine/aromatic aminotransferase	COG00436	pfam00155	cd00609	TIGR03540				2	
arCOG02165	E	Amino acid transport and metabolism	-	Transglutaminase-like cysteine protease	COG01305	pfam01841							1
arCOG01131	E	Amino acid transport and metabolism	-	Aspartate/tyrosine/aromatic aminotransferase	COG00436	pfam00155	cd00609	TIGR01265				1	
arCOG04044	E	Amino acid transport and metabolism	-	2-amino-3,7-dideoxy-D-threo-hept-6-ulosonic acid synthase, DhnA-aldolase f	COG01830	pfam01791	cd00958	TIGR01949				2	1
arCOG01888	E	Amino acid transport and metabolism	AmpS	Leucyl aminopeptidase (aminopeptidase T)	COG02309	pfam02073						2	
arCOG01890	E	Amino acid transport and metabolism	AmpS	Leucyl aminopeptidase (aminopeptidase T)	COG02309	pfam02073						2	
arCOG01924	E	Amino acid transport and metabolism	AnsB	L-asparaginase/archaeal Glu-tRNAGln amidotransferase subunit D	COG00252	pfam00710	cd08962	TIGR02153				1	
arCOG00863	E	Amino acid transport and metabolism	ArcC	Carbamate kinase	COG00549	pfam00696	cd04235	TIGR00746		1		2	
arCOG01107	E	Amino acid transport and metabolism	ArgE	Acetylornithine deacetylase/Succinyl-diaminopimelate desuccinylase or relat	COG00624	pfam01546	cd05650	TIGR01910		1	1		
arCOG00912	E	Amino acid transport and metabolism	ArgF	Ornithine carbamoyltransferase	COG00078	pfam02729,pfam00185		TIGR00658	3				1
arCOG04134	E	Amino acid transport and metabolism	AroA	5-enolpyruvylshikimate-3-phosphate synthase	COG00128	pfam00275	cd01556	TIGR01356				1	1
arCOG01033	E	Amino acid transport and metabolism	AroE	Shikimate 5-dehydrogenase	COG00169	pfam08501,pfa	cd01065	TIGR00507				1	
arCOG01025	E	Amino acid transport and metabolism	AroK2	Archaeal shikimate kinase	COG01685	pfam00288		TIGR01920				2	1
arCOG00494	E	Amino acid transport and metabolism	Asd	Aspartate-semialdehyde dehydrogenase	COG00136	pfam01118,pfam02774		TIGR00978				1	1
arCOG00121	E	Amino acid transport and metabolism	AsnB	Asparagine synthase (glutamine-hydrolyzing)	COG00367	pfam00733			1				
arCOG00071	E	Amino acid transport and metabolism	AsnB	Asparagine synthase (glutamine-hydrolyzing)	COG00367	pfam00733	cd01991	TIGR01536		1		2	1
arCOG01594	E	Amino acid transport and metabolism	CarB	Carbamoylphosphate synthase large subunit	COG00458	pfam00289,pfam02786,pfam02		TIGR01369	1				
arCOG00073	E	Amino acid transport and metabolism	CysH	PAPS reductase related enzyme fused to RNA-binding PUA domain and ferre	COG00175	pfam01507	cd01713	TIGR00434				2	
arCOG01430	E	Amino acid transport and metabolism	CysK	Cysteine synthase	COG00031	pfam00291	cd01561	TIGR01136		1	1		1
arCOG00755	E	Amino acid transport and metabolism	DadA	Glycine/D-amino acid oxidase (deaminating)	COG00665	pfam01266			2			2	
arCOG01646	E	Amino acid transport and metabolism	DAP2	Dipeptidyl aminopeptidase/acylaminoacyl-peptidase	COG01506	pfam07859	cd00312		8	1		2	
arCOG03109	E	Amino acid transport and metabolism	DdaH	N-Dimethylarginine dimethylaminohydrolase	COG01834	pfam02274		TIGR01078		1		3	
arCOG05395	E	Amino acid transport and metabolism	FtcD	Formiminotetrahydrofolate cyclodeaminase	COG03404							2	
arCOG00915	E	Amino acid transport and metabolism	GabT	4-aminobutyrate aminotransferase or related aminotransferase	COG00160	pfam00202	cd00610	TIGR00707		1			
arCOG01303	E	Amino acid transport and metabolism	GcvH	Glycine cleavage system H protein (lipoate-binding)	COG00509	pfam01597	cd06848	TIGR00527				1	
arCOG00756	E	Amino acid transport and metabolism	GcvT	Glycine cleavage system T protein (aminomethyltransferase)	COG00404	pfam01571,pfam08669		TIGR00528					1
arCOG00757	E	Amino acid transport and metabolism	GcvT	Glycine cleavage system T protein (aminomethyltransferase)	COG00404	pfam01571,pfam08669		TIGR00528				2	
arCOG02706	E	Amino acid transport and metabolism	GloA	Lactoylglutathione lyase or related enzyme	COG00346	pfam13669	cd07249	TIGR03081	1				
arCOG00095	E	Amino acid transport and metabolism	GltB	Glutamate synthase domain 1	COG00067	pfam00310	cd01907	TIGR01134					1
arCOG00096	E	Amino acid transport and metabolism	GltB	Glutamate synthase domain 3	COG00070	pfam01493	cd00981	TIGR03122		1		1	
arCOG00619	E	Amino acid transport and metabolism	GltB	Glutamate synthase domain 2 and ferredoxin domain	COG00069	pfam01645	cd02808					1	
arCOG07384	E	Amino acid transport and metabolism	HisB	Histidinol phosphatase or related phosphatase	COG00241	pfam13242	cd01427	TIGR01656					1
arCOG04273	E	Amino acid transport and metabolism	HisC	Histidinol-phosphate/aromatic aminotransferase or cobyric acid decarboxyla	COG00079	pfam00155	cd00609	TIGR01140				1	
arCOG01799	E	Amino acid transport and metabolism	HisJ	ABC-type amino acid transport/signal transduction system, periplasmic comp	COG00834	pfam00497	cd13624	TIGR01096				1	
arCOG01798	E	Amino acid transport and metabolism	HisM	ABC-type amino acid transport system, permease component	COG00765	pfam00528	cd06261	TIGR03003				1	
arCOG04671	E	Amino acid transport and metabolism	HutH	Histidine ammonia-lyase	COG02986	pfam00221	cd00332	TIGR01225				2	
arCOG04670	E	Amino acid transport and metabolism	HutU	Urocanate hydratase	COG02987	pfam01175		TIGR01228	5				
arCOG01512	E	Amino acid transport and metabolism	HyuB	N-methylhydantoinase B/acetone carboxylase, alpha subunit	COG00146	pfam02538			1				
arCOG04779	E	Amino acid transport and metabolism	IaaA	Isoaspartyl peptidase or L-asparaginase, Ntn-hydrolase superfamily	COG01446	pfam01112	cd14950					1	
arCOG02297	E	Amino acid transport and metabolism	IlvE	Branched-chain amino acid aminotransferase/4-amino-4-deoxychorismate ly	COG00115	pfam01063	cd01558	TIGR01122	3				1
arCOG01698	E	Amino acid transport and metabolism	LeuC	Homoaconitate hydratase/3-isopropylmalate dehydratase large subunit fami	COG00065	pfam00330	cd01583	TIGR02086				1	
arCOG01021	E	Amino acid transport and metabolism	LivK	ABC-type branched-chain amino acid transport system, periplasmic compone	COG00683	pfam13458	cd06268					1	
arCOG00243	E	Amino acid transport and metabolism	LYS9	Saccharopine dehydrogenase or related enzyme	COG01748	pfam03435			5	1			
arCOG00861	E	Amino acid transport and metabolism	LysC	Aspartokinase	COG00527								1
arCOG00060	E	Amino acid transport and metabolism	MetC	Cystathionine beta-lyase/cystathionine gamma-synthase	COG00626	pfam01053	cd00614	TIGR02080			1		1
arCOG00027	E	Amino acid transport and metabolism	MfnA	L-tyrosine decarboxylase, PLP-dependent protein	COG00076	pfam00282	cd06450	TIGR03812	1				1
arCOG04322	E	Amino acid transport and metabolism	PepB	Leucyl aminopeptidase	COG00260	pfam02789,pfa	cd00433					2	
arCOG04758	E	Amino acid transport and metabolism	PepF	Oligoendopeptidase F	COG01164	pfam08439,pfa	cd09608	TIGR00181				1	
arCOG01000	E	Amino acid transport and metabolism	PepP	Xaa-Pro aminopeptidase	COG00006	pfam01321,pfa	cd01092	TIGR00500			1		
arCOG00177	E	Amino acid transport and metabolism	PotA	ABC-type spermidine/putrescine transport system, ATPase component	COG03842	pfam00005	cd03301	TIGR03265				1	
arCOG06322	E	Amino acid transport and metabolism	PutA	Proline dehydrogenase	COG00506	pfam01619			3				1
arCOG01316	E	Amino acid transport and metabolism	PutP	Na+/proline symporter	COG00591	pfam00474	cd10322	TIGR00813	1				
arCOG00088	E	Amino acid transport and metabolism	PuuD	Gamma-glutamyl-gamma-aminobutyrate hydrolase PuuD (putrescine degrad	COG02071	pfam07722	cd01745	TIGR00888				4	
arCOG01947	E	Amino acid transport and metabolism	RhtB	Threonine efflux protein	COG01280	pfam01810			3				
arCOG01948	E	Amino acid transport and metabolism	RhtB	Threonine efflux protein	COG01280	pfam01810		TIGR00949					1
arCOG01158	E	Amino acid transport and metabolism	SerB	Phosphoserine phosphatase	COG00560	pfam00702		TIGR01491	5	1		1	2
arCOG01459	E	Amino acid transport and metabolism	Tdh	Threonine dehydrogenase or related Zn-dependent dehydrogenase	COG01063	pfam08240,pfa	cd08235	TIGR00692					1
arCOG05121	E	Amino acid transport and metabolism	TesA	Lysophospholipase L1 or related esterase	COG02755				1				
arCOG01434	E	Amino acid transport and metabolism	ThrC	Threonine synthase and cysteate synthase	COG00498	pfam01155,pfa	cd01563	TIGR02605,TIGR00260				1	1
arCOG02014	E	Amino acid transport and metabolism	TrpE	Anthranilate/para-aminobenzoate synthase component I	COG00147	pfam04715,pfam00425		TIGR01820				1	1
arCOG07581	G	Carbohydrate transport and metabolism	-	Glycosyl hydrolase family 18, contains cellulose binding domain	COG03325	pfam02839,pfa	cd12215,cd00146,cd06548		7		1	3	3
arCOG00138	G	Carbohydrate transport and metabolism	-	MFS family permease	COG00477				5				
arCOG03641	G	Carbohydrate transport and metabolism	PfkA	6-phosphofructokinase	COG00205	pfam00365	cd00363	TIGR02483	5				
arCOG04431	G	Carbohydrate transport and metabolism	GlpF	Glycerol uptake facilitator or related permease (Major Intrinsic Protein Famil	COG00580	pfam00230	cd00333	TIGR00861	4	1		1	
arCOG00130	G	Carbohydrate transport and metabolism	-	MFS family permease	COG00477	pfam07690	cd06174		3	1			
arCOG01349	G	Carbohydrate transport and metabolism	SuhB	Archaeal fructose-1,6-bisphosphatase or related enzyme of inositol monoph	COG00483	pfam00459	cd01642	TIGR01331	2	1	1	1	
arCOG00271	G	Carbohydrate transport and metabolism	RhaT	Permease of the drug/metabolite transporter (DMT) superfamily	COG00697	pfam00892,pfam00892		TIGR00950	2				1
arCOG02876	G	Carbohydrate transport and metabolism	CDA1	Peptidoglycan/xylan/chitin deacetylase, PgdA/CDA1 family	COG00726	pfam01522	cd10941	TIGR03006	2				
arCOG01122	G	Carbohydrate transport and metabolism	RpiA	Ribose 5-phosphate isomerase	COG00120	pfam06026	cd01398	TIGR00021	1				
arCOG00052	G	Carbohydrate transport and metabolism	Pgi	Glucose-6-phosphate isomerase	COG00166	pfam01380,pfa	cd05017,cd056	TIGR02128	1	1		1	1
arCOG01111	G	Carbohydrate transport and metabolism	PpsA	Phosphoenolpyruvate synthase/pyruvate phosphate dikinase	COG00574	pfam01326,pfam00391,pfam02		TIGR01418	1	1			
arCOG04934	G	Carbohydrate transport and metabolism	-	Peptidoglycan/xylan/chitin deacetylase, PgdA/CDA1 family	COG00726				1				
arCOG07840	G	Carbohydrate transport and metabolism	-	Glycosyl hydrolase family 18, contains cellulose binding domain	COG03325	pfam02839,pfa	cd12215,cd00146,cd06543		1			2	
arCOG00767	G	Carbohydrate transport and metabolism	(ManB)	Phosphomannomutase	COG01109	pfam02878,pfa	cd03087	TIGR03990					1
arCOG00274	G	Carbohydrate transport and metabolism	RhaT	Permease of the drug/metabolite transporter (DMT) superfamily	COG00697	pfam00892,pfam00892		TIGR00950		1	1	2	
arCOG00014	G	Carbohydrate transport and metabolism	RbsK	Sugar kinase, ribokinase family	COG00524	pfam00294	cd01174	TIGR02152		1		5	
arCOG05046	G	Carbohydrate transport and metabolism	Rpe	Pentose-5-phosphate-3-epimerase	COG00036	pfam00834	cd00429	TIGR01163				1	
arCOG00496	G	Carbohydrate transport and metabolism	Pgk	3-phosphoglycerate kinase	COG00126	pfam00162	cd00318					1	1
arCOG02796	G	Carbohydrate transport and metabolism	-	Glucose/sorbosone dehydrogenase	COG02133	pfam07995		TIGR03606			1	1	1
arCOG00135	G	Carbohydrate transport and metabolism	-	MFS family permease	COG00477	pfam05977	cd06174	TIGR00900				2	
arCOG01029	G	Carbohydrate transport and metabolism	GalK	Galactokinase	COG00153	pfam10509,pfam08544		TIGR00131					1
arCOG01393	G	Carbohydrate transport and metabolism	-	Glycosyl transferase, related to UDP-glucuronosyltransferase	COG01819								1

arCOG04120	G	Carbohydrate transport and metabolism	PykF	Pyruvate kinase	COG00469	pfam00224,pfa	cd00288	TIGR01064					1
arCOG05061	G	Carbohydrate transport and metabolism	TalA	Transaldolase	COG00176	pfam00923	cd00956	TIGR00875					1
arCOG05412	G	Carbohydrate transport and metabolism	BglB	Beta-glucosidase/6-phospho-beta-glucosidase/beta-galactosidase	COG02723	pfam00232		TIGR03356					1
arCOG04339	G	Carbohydrate transport and metabolism	TagG	ABC-type polysaccharide/polyol phosphate export systems, permease compc	COG01682	pfam01061						1	1
arCOG04263	H	Coenzyme transport and metabolism	-	Pantoate kinase	COG01829	pfam00288			4			1	
arCOG00021	H	Coenzyme transport and metabolism	-	Predicted transcriptional regulator fused phosphomethylpyrimidine kinase, i	COG01992	pfam10120			4				1
arCOG01484	H	Coenzyme transport and metabolism	RibD	Pyrimidine reductase, riboflavin biosynthesis	COG01985	pfam01872		TIGR01508	3	1	1		
arCOG01940	H	Coenzyme transport and metabolism	BirA	Biotin-(acetyl-CoA carboxylase) ligase	COG00340	pfam03099		TIGR00121	3	1	1	2	
arCOG00040	H	Coenzyme transport and metabolism	Hpt1	Hypoxanthine phosphoribosyltransferase	COG02236	pfam00156	cd06223	TIGR01251	3	1			
arCOG02172	H	Coenzyme transport and metabolism	QueD	6-pyruvoyl-tetrahydropterin synthase	COG00720	pfam01242	cd00470	TIGR03367	3		1		
arCOG01045	H	Coenzyme transport and metabolism	CoaE	Dephospho-CoA kinase	COG00237		cd02022		3		1	1	
arCOG00226	H	Coenzyme transport and metabolism	TbpA	ABC-type thiamine transport system, periplasmic component	COG04143	pfam13343	cd13545	TIGR01254	2				1
arCOG01589	H	Coenzyme transport and metabolism	RimK	Glutathione synthase/glutaminyl transferase/alpha-L-glutamate ligase	COG00189	pfam08443		TIGR02144	2				
arCOG03402	H	Coenzyme transport and metabolism	MtbC1	Methanogenic corrinoid protein MtbC1	COG05012	pfam02607,pfa	cd02065	TIGR02370	2				
arCOG00476	H	Coenzyme transport and metabolism	UbiA	4-hydroxybenzoate polyprenyltransferase or related prenyltransferase	COG00382	pfam01040	cd13961	TIGR01476	1				1
arCOG00972	H	Coenzyme transport and metabolism	NadR	Nicotinamide mononucleotide adenylyltransferase	COG01056	pfam01467	cd02166	TIGR01527	1	1	1	1	1
arCOG04139	H	Coenzyme transport and metabolism	ApbA	Ketopantoate reductase	COG01893	pfam02558,pfam08546		TIGR00745	1	1	1		
arCOG01522	H	Coenzyme transport and metabolism	HemY	Protoporphyrinogen oxidase	COG01232	pfam01593		TIGR00562	1				
arCOG01490	H	Coenzyme transport and metabolism	FolA	Dihydrofolate reductase	COG00262	pfam00186	cd00209					1	
arCOG01223	H	Coenzyme transport and metabolism	CAB4	Phosphopantetheine adenylyltransferase	COG01019	pfam01467	cd02164	TIGR00125		1	1	1	
arCOG01320	H	Coenzyme transport and metabolism	RibB	3,4-dihydroxy-2-butanone 4-phosphate synthase	COG00108	pfam00926		TIGR00506		1			
arCOG01671	H	Coenzyme transport and metabolism	UbiD	3-polyprenyl-4-hydroxybenzoate decarboxylase or related decarboxylase	COG00043	pfam01977		TIGR03701		1			
arCOG02939	H	Coenzyme transport and metabolism	PhhB	Pterin-4a-carbinolamine dehydratase	COG02154	pfam01329	cd00488			1		1	
arCOG01482	H	Coenzyme transport and metabolism	NadC	Nicotinate-nucleotide pyrophosphorylase	COG00157	pfam02749,pfa	cd01572	TIGR00078				1	
arCOG04137	H	Coenzyme transport and metabolism	SAM1	S-adenosylhomocysteine hydrolase	COG00499	pfam05221	cd00401	TIGR00936		1			
arCOG03838	H	Coenzyme transport and metabolism	PqqD	Coenzyme PQQ synthesis protein D		pfam05402				1	1		1
arCOG00069	H	Coenzyme transport and metabolism	NadE	NH3-dependent NAD+ synthetase	COG00171	pfam02540	cd00553	TIGR00552				3	1
arCOG01942	H	Coenzyme transport and metabolism	LipB	Lipoate-protein ligase B	COG00321	pfam03099		TIGR00214		1			
arCOG00480	H	Coenzyme transport and metabolism	MenA	1,4-dihydroxy-2-naphthoate octaprenyltransferase	COG01575	pfam01040	cd13962	TIGR02235			1		
arCOG00489	H	Coenzyme transport and metabolism	PduO	Cob(I)alamin adenosyltransferase	COG02096	pfam01923		TIGR00636					1
arCOG02817	H	Coenzyme transport and metabolism	FolC	Folylpolyglutamate synthase and Dihydropteroate synthase	COG00285	pfam08245,pfa	cd00739	TIGR01499,TIGR01496				1	1
arCOG01704	H	Coenzyme transport and metabolism	Dfp	Phosphopantothenoylcysteine synthetase/decarboxylase	COG00452	pfam02441,pfam04127		TIGR00521				2	1
arCOG00214	H	Coenzyme transport and metabolism	MoaB	Molybdopterin biosynthesis enzyme	COG00521	pfam00994	cd00886	TIGR02667			1		
arCOG00534	H	Coenzyme transport and metabolism	MoaE	Molybdopterin converting factor, large subunit	COG00314	pfam02391	cd00756				1		
arCOG00930	H	Coenzyme transport and metabolism	MoaA	Molybdenum cofactor biosynthesis enzyme	COG02896	pfam04055,pfa	cd01335	TIGR02668			1	1	
arCOG01530	H	Coenzyme transport and metabolism	MoaC	Molybdenum cofactor biosynthesis enzyme	COG00315	pfam01967	cd01419	TIGR00581			1		
arCOG01872	H	Coenzyme transport and metabolism	MobA	Molybdopterin-guanine dinucleotide biosynthesis protein A	COG00746	pfam12804	cd02503				1	2	
arCOG00217	H	Coenzyme transport and metabolism	MoeA	Molybdopterin biosynthesis enzyme	COG00303	pfam03453,pfa	cd00887	TIGR00177			1	3	
arCOG02199	H	Coenzyme transport and metabolism	Mcr1	NAD(P)H-flavin reductase	COG00543						1		
arCOG01676	H	Coenzyme transport and metabolism	ThiF	Dinucleotide-utilizing enzyme involved in molybdopterin and thiamine biosyr	COG00476	pfam00899,pfa	cd00757	TIGR02356				2	1
arCOG01873	H	Coenzyme transport and metabolism	MobA	GT-A family glycosyltransferase involved in molybdopterin guanine dinucleot	COG02068	pfam12804	cd04182	TIGR03310				1	1
arCOG04348	H	Coenzyme transport and metabolism	UbiE	Ubiquinone/menaquinone biosynthesis C-methylase UbiE	COG02226	pfam08241		TIGR01934				1	1
arCOG00572	H	Coenzyme transport and metabolism	NadB	Aspartate oxidase	COG00029	pfam00890,pfam02910		TIGR00551				1	
arCOG02624	H	Coenzyme transport and metabolism	PaaK	Coenzyme F390 synthetase	COG01541			TIGR02304					1
arCOG00570	C	Energy production and conversion	FixC	Dehydrogenase (flavoprotein)	COG00644	pfam13450		TIGR02032	6	2		2	1
arCOG02921	C	Energy production and conversion	PetE	Plastocyanin	COG03794	pfam00127	cd04220	TIGR03102	4				
arCOG00701	C	Energy production and conversion	UgpQ	Glycerophosphoryl diester phosphodiesterase	COG00584	pfam03009	cd08568		4				
arCOG00519	C	Energy production and conversion	FldA	Flavodoxin	COG00716	pfam12682			3			1	
arCOG02017	C	Energy production and conversion	RutF	NADH-FMN oxidoreductase RutF, flavin reductase (DIM6/NTAB) family	COG01853	pfam01613		TIGR03615	2			1	1
arCOG02929	C	Energy production and conversion	PetE	Plastocyanin	COG03794	pfam00127	cd13921	TIGR02657	1				
arCOG02398	C	Energy production and conversion	CcdA	Cytochrome c biogenesis protein	COG00785	pfam02683			1		1		
arCOG01599	C	Energy production and conversion	PorB	Pyruvate:ferredoxin oxidoreductase or related 2-oxoacid:ferredoxin oxidores	COG01013	pfam02775,pfa	cd03375	TIGR02177	1				
arCOG00571	C	Energy production and conversion	SdhA	Succinate dehydrogenase/fumarate reductase, flavoprotein subunit	COG01053	pfam00890,pfam02910		TIGR03378,TIGR01812	1			4	1
arCOG01551	C	Energy production and conversion	NuoC	NADH dehydrogenase subunit C	COG00852	pfam00329		TIGR01961				3	1
arCOG01539	C	Energy production and conversion	NuoL	NADH dehydrogenase subunit L	COG01009	pfam00662,pfam00361		TIGR01974					1
arCOG01540	C	Energy production and conversion	NuoN	NADH dehydrogenase subunit N	COG01007	pfam00361		TIGR01770		1			
arCOG02095	C	Energy production and conversion	OadA1	Pyruvate/oxaloacetate carboxyltransferase	COG05016	pfam00682,pfa	cd07937	TIGR01108			1	1	1
arCOG04138	C	Energy production and conversion	NtpI	Archaeal/vacuolar-type H+-ATPase subunit I	COG01269	pfam01496						1	
arCOG00869	C	Energy production and conversion	NtpE	Archaeal/vacuolar-type H+-ATPase subunit E	COG01390					1	1		
arCOG01054	C	Energy production and conversion	AcoA	Pyruvate/2-oxoglutarate/acetoin dehydrogenase complex, dehydrogenase (E	COG01071	pfam00676	cd02000	TIGR03181				1	
arCOG04279	C	Energy production and conversion	-	Swiveling domain associated with predicted aconitase	COG01786	pfam01989	cd01356			1			
arCOG00573	C	Energy production and conversion	SdhA	Succinate dehydrogenase/fumarate reductase, flavoprotein subunit	COG01053	pfam00890,pfam02910		TIGR02061		1			
arCOG02618	C	Energy production and conversion	-	Ferredoxin	COG01146	pfam13187,pfam12139		TIGR02060		1			
arCOG01235	C	Energy production and conversion	CyoA	Heme/copper-type cytochrome/quinol oxidase, subunit 2	COG01622	pfam00116	cd13918	TIGR02866		1	1	1	
arCOG00446	C	Energy production and conversion	FixA	Electron transfer flavoprotein, beta subunit	COG02086	pfam01012	cd01714					1	
arCOG01252	C	Energy production and conversion	PutA	Lactaldehyde dehydrogenase, Succinate semialdehyde dehydrogenase or oth	COG01012	pfam00171	cd07088	TIGR01780		1		3	
arCOG00288	C	Energy production and conversion	NfnB	Nitroreductase	COG00778	pfam00881	cd02150	TIGR02476				1	
arCOG01458	C	Energy production and conversion	Qor	NADPH:quinone reductase or related Zn-dependent oxidoreductase	COG00604	pfam08240,pfa	cd08264	TIGR02824				1	1
arCOG00963	C	Energy production and conversion	GlpC	Succinate dehydrogenase/fumarate reductase, Fe-S protein subunit	COG00479	pfam13085,pfam13183,pfam02		TIGR00384,TIGR03288				2	
arCOG04650	C	Energy production and conversion	CyoC	Heme/copper-type cytochrome/quinol oxidase, subunit 3	COG01845	pfam00510	cd00386	TIGR02842					2
arCOG00338	C	Energy production and conversion	SdhC	Succinate dehydrogenase subunit C	COG02048	pfam02754,pfam02754		TIGR03288				2	1
arCOG03363	C	Energy production and conversion	NtpF	Archaeal/vacuolar-type H+-ATPase subunit H	COG02811							1	
arCOG00962	C	Energy production and conversion	FrdB	Succinate dehydrogenase/fumarate reductase, Fe-S protein subunit	COG00479	pfam13085,pfam13183		TIGR00384				2	1
arCOG01617	C	Energy production and conversion	Tas	Aryl-alcohol dehydrogenase related enzyme	COG00667	pfam00248	cd06660	TIGR01293				3	1
arCOG04358	C	Energy production and conversion	FdhD	Uncharacterized protein required for formate dehydrogenase activity	COG01526	pfam02634		TIGR00129				2	
arCOG01236	C	Energy production and conversion	CyoA	Heme/copper-type cytochrome/quinol oxidase, subunit 2	COG01622	pfam00116	cd13842	TIGR02866					1
arCOG02304	C	Energy production and conversion	Mct/CaiB	Succinyl-CoA:mesaconate CoA-transferase or predicted acyl-CoA transferase,	COG01804	pfam02515		TIGR03253					1
arCOG02476	C	Energy production and conversion	HdrA	Heterodisulfide reductase, subunit A; ferredoxin domain	COG01148	pfam07992,pfam13237,pfam1		TIGR03140,TIGR04105,TIGR03385,TIGR04105					1
arCOG02842	C	Energy production and conversion	Fdx	Ferredoxin	COG00633	pfam00111	cd00207	TIGR02008				1	
arCOG09746	P	Inorganic ion transport and metabolism	-	Sulfotransferase related protein		pfam00685			4		1		
arCOG04750	P	Inorganic ion transport and metabolism	-	Sirohydrochlorin iron chelatase fused to [2Fe-2S] Ferredoxin	COG02138	pfam01903,pfa	cd03416,cd03414,cd02980		4			1	
arCOG02881	P	Inorganic ion transport and metabolism	ECM27	Ca2+/Na+ antiporter	COG00530	pfam01699,pfam01699		TIGR00367	3	1		1	
arCOG00206	P	Inorganic ion transport and metabolism	PhnC	ABC-type phosphate/phosphonate transport system, ATPase component	COG03638	pfam00005	cd03256	TIGR02315	3				
arCOG02021	P	Inorganic ion transport and metabolism	PspE	Rhodanese-related sulfurtransferase	COG00607	pfam00581	cd00158	TIGR02981	2	1	1		3
arCOG04233	P	Inorganic ion transport and metabolism	FepB	ABC-type Fe3+-hydroxamate transport system, periplasmic component	COG00614	pfam01497	cd01143	TIGR04281	2				1
arCOG00163	P	Inorganic ion transport and metabolism	ThiP	ABC-type Fe3+ transport system, permease component	COG01178	pfam00528,pfa	cd06261,cd062	TIGR03262	2				1
arCOG02763	P	Inorganic ion transport and metabolism	-	Heavy-metal-associated domain (HMA)	COG02608	pfam00403	cd00371	TIGR02052	1			1	
arCOG00238	P	Inorganic ion transport and metabolism	ArsB	Na+/H+ antiporter NhaD or related arsenite permease	COG01055	pfam03600	cd01117	TIGR00785	1				
arCOG01005	P	Inorganic ion transport and metabolism	LraI	ABC-type metal ion transport system, periplasmic component/surface adhes	COG00803	pfam01297	cd01018	TIGR03772,TIGR03772	1				
arCOG02851	P	Inorganic ion transport and metabolism	(NirD)	Ferredoxin subunit of nitrite reductase or ring-hydroxylating dioxygenase	COG02146	pfam00355	cd03467		1				
arCOG08934	P	Inorganic ion transport and metabolism	-	Alkylhydroperoxidase family enzyme	COG02128	pfam02627			1				
arCOG04487	P	Inorganic ion transport and metabolism	KatG	Catalase (peroxidase I)	COG00376	pfam00141,pfa	cd00649,cd006	TIGR00198	1				
arCOG02497	P	Inorganic ion transport and metabolism	-	Lipoprotein NosD family, contains CASH domains	COG03420	pfam12708		TIGR04247	1				
arCOG00318	P	Inorganic ion transport and metabolism	PhoU	Phosphate uptake regulator	COG00704	pfam04014,pfam01895,pfam0		TIGR02135	1			1	1
arCOG06837	P	Inorganic ion transport and metabolism	-	Lipoprotein NosD family, contains CASH domains	COG03420	pfam05048		TIGR04247,TIGR03024	1				2
arCOG01959	P	Inorganic ion transport and metabolism	TrkA	TrkA, K+ transport system, NAD-binding component	COG00569	pfam02254,pfam02080,pfam02254,pfam02080						1	1
arCOG04231	P	Inorganic ion transport and metabolism	CutA	Uncharacterized protein involved in tolerance to divalent cations	COG01324	pfam03091						3	1
arCOG02499	P	Inorganic ion transport and metabolism	-	Lipoprotein NosD family, contains CASH domains	COG03420	pfam13229,pfam13229		TIGR04247			1		
arCOG02265	P	Inorganic ion transport and metabolism	CorA	Mg2+ and Co2+ transporter	COG00598	pfam01544	cd12828	TIGR00383			1		
arCOG04397	P	Inorganic ion transport and metabolism	AmtB	Ammonia permease	COG00004	pfam00909		TIGR00836			1		
arCOG02764	P	Inorganic ion transport and metabolism	CopZ	Copper-ion-binding protein	COG02608	pfam00403	cd00371	TIGR02052					1
arCOG01805	P	Inorganic ion transport and metabolism	PhnD	ABC-type phosphate/phosphonate transport system, periplasmic component	COG03221	pfam12974	cd01071	TIGR01098				1	
arCOG02050	P	Inorganic ion transport and metabolism	-	Sulfite exporter, TauE/SafE family	COG00730	pfam01925						2	1
arCOG00202	P	Inorganic ion transport and metabolism	EcfA2	Energy-coupling factor transporter ATP-binding protein EcfA2	COG01122	pfam00005	cd03225	TIGR01166					1
arCOG01040	P	Inorganic ion transport and metabolism	CysC	Adenylylsulfate kinase or related kinase	COG00529	pfam01583	cd02027	TIGR00455				1	
arCOG01096	P	Inorganic ion transport and metabolism	CCC1	Predicted Fe2+/Mn2+ transporter, VIT1/CCC1 family	COG01814	pfam02915,pfa	cd01044,cd02431					1	
arCOG01576	P	Inorganic ion transport and metabolism	ZntA	Cation transport ATPase	COG02217	pfam00403,pfa	cd00371,cd014	TIGR00003,TIGR01511				1	1
arCOG02062	P	Inorganic ion transport and metabolism	TusA	TusA-related sulfurtransferase	COG00425	pfam01206	cd00291					1	1
arCOG05356	P	Inorganic ion transport and metabolism	Fes	Enterochelin esterase or related enzyme	COG02382	pfam00756						1	1
arCOG11907	P	Inorganic ion transport and metabolism	-	Fe-S metabolizm associated domain		pfam02657		TIGR03391				1	1
arCOG00230	P	Inorganic ion transport and metabolism	-	Periplasmic molybdate-binding protein/domain	COG01910	pfam00126,pfam12727		TIGR00637					1
arCOG02849	P	Inorganic ion transport and metabolism	ArsA	Oxyanion-translocating ATPase	COG00003	pfam02374	cd02035	TIGR00345					1
arCOG04559	P	Inorganic ion transport and metabolism	EmrE	Membrane transporter of cations and cationic drugs	COG02076	pfam00893							1
arCOG01964	P	Inorganic ion transport and metabolism	Kch	Kef-type K+ transport system, predicted NAD-binding component	COG01226	pfam07885							2
arCOG00167	P	Inorganic ion transport and metabolism	PstC	ABC-type phosphate transport system, permease component	COG00573		cd06261	TIGR02138				1	
arCOG00168	P	Inorganic ion transport and metabolism	PstA	ABC-type phosphate transport system, permease component	COG00581		cd06261	TIGR00974				1	
arCOG00213	P	Inorganic ion transport and metabolism	PstS	ABC-type phosphate transport system, periplasmic component	COG00226	pfam12849	cd13565	TIGR00975				1	
arCOG00231	P	Inorganic ion transport and metabolism	PstB	ABC-type phosphate transport system, ATPase component	COG01117	pfam00005	cd03260	TIGR00972				1	
arCOG00232	P	Inorganic ion transport and metabolism	PhoU	Phosphate uptake regulator	COG00704			TIGR02135				1	
arCOG02555	P	Inorganic ion transport and metabolism	-	Lipoprotein NosD family, contains CASH domains	COG03420							1	
arCOG02852	P	Inorganic ion transport and metabolism	(NirD)	Ferredoxin subunit of nitrite reductase or ring-hydroxylating dioxygenase	COG02146	pfam00355	cd03528	TIGR02377				1	
arCOG03797	P	Inorganic ion transport and metabolism	-	Ferritin-like domain	COG01633	pfam05763						1	
arCOG01259	I	Lipid transport and metabolism	FabG	Short-chain alcohol dehydrogenase	COG01028		cd05233	TIGR01830	5			3	5
arCOG00670	I	Lipid transport and metabolism	PgsA	Phosphatidylglycerophosphate synthase	COG00558	pfam01066			5	1			2
arCOG01767	I	Lipid transport and metabolism	PksG	3-hydroxy-3-methylglutaryl CoA synthase	COG03425	pfam08545,pfa	cd00827	TIGR00748	5			1	2
arCOG00671	I	Lipid transport and metabolism	PssA	Phosphatidylserine synthase	COG01183	pfam01066		TIGR04217	5				
arCOG13950	I	Lipid transport and metabolism	-	Phosphatidylglycerophosphatase A		pfam04608	cd06971		4				
arCOG04199	I	Lipid transport and metabolism	FAA1	Long-chain acyl-CoA synthetase (AMP-forming)	COG01022	pfam00501	cd05907	TIGR01923	4				1
arCOG01278	I	Lipid transport and metabolism	PaaJ	Acetyl-CoA acetyltransferase	COG00183		cd00829	TIGR01930	4				1
arCOG08932	I	Lipid transport and metabolism	LCB5	Diacylglycerol kinase family enzyme	COG01597	pfam00781		TIGR03702	3	1	1	2	1
arCOG00860	I	Lipid transport and metabolism	-	Isopentenyl phosphate kinase, enzyme of modified mevalonate pathway	COG01608	pfam00696	cd04241	TIGR02075	3	1		1	1
arCOG01028	I	Lipid transport and metabolism	ERG12	Mevalonate kinase	COG01577	pfam00288		TIGR00549	3	1		3	1
arCOG02936	I	Lipid transport and metabolism	ERG9	Phytoene/squalene synthetase	COG01562	pfam00494	cd00683	TIGR03465	3				
arCOG01137	I	Lipid transport and metabolism	FadM	Acyl-CoA thioesterase FadM	COG00824	pfam13279	cd00586	TIGR00051	3		1		
arCOG01879	I	Lipid transport and metabolism	-	Dolichol kinase family protein	COG00170				3				1
arCOG02245	I	Lipid transport and metabolism	-	Cytidylyltransferase family enzyme	COG01836	pfam01940		TIGR00297	2			2	1
arCOG04351	I	Lipid transport and metabolism	-	Predicted membrane associated lipid hydrolase, neutral ceramidase superfar	COG03356	pfam09843			2	1			
arCOG00249	I	Lipid transport and metabolism	FadB	3-hydroxyacyl-CoA dehydrogenase, some fused to Enoyl-CoA hydratase	COG01250	pfam02737,pfa	cd06558	TIGR02437,TIGR02437	1			1	
arCOG00239	I	Lipid transport and metabolism	CaiD	Enoyl-CoA hydratase/carnithine racemase	COG01024	pfam00378	cd06558	TIGR03210	1			3	
arCOG01880	I	Lipid transport and metabolism	SEC59	Dolichol kinase	COG00170				1				
arCOG00856	I	Lipid transport and metabolism	CaiC	Acyl-CoA synthetase (AMP-forming)/AMP-acid ligase II	COG00318	pfam00501	cd12119	TIGR01923	1				
arCOG09737	I	Lipid transport and metabolism	LicD	Phosphorylcholine metabolism protein LicD	COG03475	pfam04991			1				

arCOG04106	I	Lipid transport and metabolism	CdsA	CDP-diglyceride synthetase	COG00575	pfam01864						1		
arCOG01532	I	Lipid transport and metabolism	UppS	Undecaprenyl pyrophosphate synthase	COG00020	pfam01255	cd00475	TIGR00055				1		
arCOG01843	I	Lipid transport and metabolism	-	Sterol carrier protein	COG03255	pfam02036				1				
arCOG01085	I	Lipid transport and metabolism	PcrB	(S)-3-O-geranylgeranylglyceryl phosphate synthase, TIM-barrel fold	COG01646	pfam01884	cd02812	TIGR01768				2	1	
arCOG04260	I	Lipid transport and metabolism	HMG1	Hydroxymethylglutaryl-CoA reductase	COG01257	pfam00368	cd00643	TIGR00533					1	
arCOG01710	I	Lipid transport and metabolism	Sbm	Methylmalonyl-CoA mutase, C-terminal domain/subunit (cobalamin-binding)	COG02185	pfam02310	cd02071	TIGR00640				2		
arCOG02705	I	Lipid transport and metabolism	-	Acetyl-CoA carboxylase, carboxyltransferase component	COG04799	pfam01039		TIGR01117					1	
arCOG01707	I	Lipid transport and metabolism	CaiA	Acyl-CoA dehydrogenase	COG01960	pfam02771,pfa	cd00567	TIGR03207				1		
arCOG01282	I	Lipid transport and metabolism	PaaJ	Acetyl-CoA acetyltransferase	COG00183	pfam00108,pfa	cd00751	TIGR01930					1	
arCOG04470	I	Lipid transport and metabolism	Psd	Phosphatidylserine decarboxylase	COG00688	pfam02666		TIGR00164					1	
arCOG01650	I	Lipid transport and metabolism	PldB	Lysophospholipase, alpha-beta hydrolase superfamily	COG02267	pfam12697		TIGR03695				1		
arCOG02039	I	Lipid transport and metabolism	Cls	Phosphatidylserine/phosphatidylglycerophosphate/cardiolipin synthase or re	COG01502	pfam13091,pfa	cd07493,cd09127,cd09128						1	
arCOG00241	I	Lipid transport and metabolism	CaiD	Enoyl-CoA hydratase/carnithine racemase	COG01024	pfam00378	cd06558	TIGR02280				2		
arCOG01261	I	Lipid transport and metabolism	FabG	Short-chain alcohol dehydrogenase	COG01028	pfam00106	cd05233	TIGR01830				1	2	
arCOG03951	I	Lipid transport and metabolism	PgpB	Membrane-associated phospholipid phosphatase	COG00671	pfam01569	cd03386					3		
arCOG00674	I	Lipid transport and metabolism	PgsA	Phosphatidylglycerophosphate synthase	COG00558	pfam01066						1		
arCOG06122	I	Lipid transport and metabolism	Acs	Acyl-coenzyme A synthetase/AMP-(fatty) acid ligase	COG00365	pfam00501	cd05959	TIGR02262				1		
arCOG00911	F	Nucleotide transport and metabolism	PyrB	Aspartate carbamoyltransferase, catalytic chain	COG00540	pfam02729,pfam00185		TIGR00670	3					
arCOG04415	F	Nucleotide transport and metabolism	PurD	Phosphoribosylamine-glycine ligase	COG00151	pfam02844,pfam01071,pfam02		TIGR00877	3	1				
arCOG01038	F	Nucleotide transport and metabolism	Fap7	Broad-specificity NMP kinase	COG01936	pfam13238			3			1	1	
arCOG00018	F	Nucleotide transport and metabolism	Nnr2	NAD(P)H-hydrate repair enzyme Nnr, NAD(P)H-hydrate dehydratase domain	COG00063	pfam03853,pfa	cd01171	TIGR00197,TIGR00196	3	1				
arCOG01891	F	Nucleotide transport and metabolism	Tmk	Thymidylate kinase	COG00125	pfam02223	cd01672	TIGR00041	3	1		1		
arCOG00081	F	Nucleotide transport and metabolism	PyrF	Orotidine-5'-phosphate decarboxylase	COG00284	pfam00215	cd04725	TIGR01740	3					
arCOG03575	F	Nucleotide transport and metabolism	-	Polyphosphate kinase 2	COG02326	pfam03976,pfa	cd01672	TIGR03708	3					
arCOG00087	F	Nucleotide transport and metabolism	GuaA	GMP synthase - Glutamine amidotransferase domain	COG00518	pfam00117	cd01742	TIGR00888	2					
arCOG00093	F	Nucleotide transport and metabolism	PurF	Glutamine phosphoribosylpyrophosphate amidotransferase	COG00034	pfam00310,pfa	cd00715,cd062	TIGR01134	2					
arCOG04276	F	Nucleotide transport and metabolism	NrdA	Ribonucleotide reductase, alpha subunit	COG00209	pfam03477,pfa	cd02888	TIGR02504	1					
arCOG00090	F	Nucleotide transport and metabolism	GuaA	GMP synthase - Glutamine amidotransferase domain	COG00518	pfam00117	cd01741	TIGR00888	1	1		3		
arCOG02825	F	Nucleotide transport and metabolism	PurN	Folate-dependent phosphoribosylglycinamide formyltransferase PurN	COG00299				1				2	
arCOG04184	F	Nucleotide transport and metabolism	RdgB	Inosine/xanthosine triphosphate pyrophosphatase, all-alpha NTP-PPase fami	COG00127	pfam01725	cd00515	TIGR00042			1	2		
arCOG00028	F	Nucleotide transport and metabolism	-	Orotate phosphoribosyltransferase homolog	COG00856	pfam00156	cd06223	TIGR02985,TIGR00336				1	1	
arCOG00029	F	Nucleotide transport and metabolism	PyrE	Orotate phosphoribosyltransferase	COG00461	pfam00156	cd06223	TIGR00336		1				
arCOG00419	F	Nucleotide transport and metabolism	Hit	HIT family hydrolase	COG00537	pfam01230	cd01275					3	1	
arCOG04462	F	Nucleotide transport and metabolism	PurS	Phosphoribosylformylglycinamidine (FGAM) synthase, PurS component	COG01828	pfam02700		TIGR00302				1		
arCOG04421	F	Nucleotide transport and metabolism	PurC	Phosphoribosylaminoimidazolesuccinocarboxamide (SAICAR) synthase	COG00152	pfam01259	cd01415	TIGR00081					1	
arCOG00641	F	Nucleotide transport and metabolism	PurL	Phosphoribosylformylglycinamidine (FGAM) synthase, synthetase domain	COG00046	pfam00586,pfa	cd02203,cd022	TIGR01736				1		
arCOG01565	F	Nucleotide transport and metabolism	NrnA	nanoRNase/pAp phosphatase, hydrolyzes c-di-AMP and oligoRNAs	COG00618	pfam01368				1			1	
arCOG04346	F	Nucleotide transport and metabolism	-	5-formaminoimidazole-4-carboxamide-1-beta-D-ribofuranosyl 5'-monophosp	COG01759	pfam06849,pfam06973		TIGR00877		1			2	
arCOG00603	F	Nucleotide transport and metabolism	PyrD	Dihydroorotate dehydrogenase	COG00167	pfam01180	cd04740	TIGR01037			1	2		
arCOG01037	F	Nucleotide transport and metabolism	Cmk	Cytidylate kinase	COG01102	pfam13189	cd02020	TIGR02173		1		1	1	
arCOG00612	F	Nucleotide transport and metabolism	GuaB	IMP dehydrogenase/GMP reductase	COG00516	pfam00478	cd00381	TIGR01302				2		
arCOG05252	F	Nucleotide transport and metabolism	-	Predicted secreted endonuclease distantly related to archaeal Holliday junct	COG04741	pfam10107					1			
arCOG00030	F	Nucleotide transport and metabolism	Apt	Adenine/guanine phosphoribosyltransferase or related PRPP-binding protein	COG00503	pfam00156	cd06223	TIGR01090				1		
arCOG00858	F	Nucleotide transport and metabolism	PyrH	Uridylate kinase	COG00528	pfam00696	cd04253	TIGR02076				1	1	
arCOG02013	F	Nucleotide transport and metabolism	DeoA	Thymidine phosphorylase	COG00213	pfam01568,pfam02885,pfam0		TIGR03327					1	
arCOG04311	F	Nucleotide transport and metabolism	-	5'-deoxynucleotidase, HD superfamily hydrolase	COG01896	pfam13023							1	
arCOG01075	F	Nucleotide transport and metabolism	-	NUDIX family hydrolase	COG01051	pfam00293	cd04673					7	3	
arCOG04320	F	Nucleotide transport and metabolism	DeoC	Deoxyribose-phosphate aldolase	COG00274	pfam01791	cd00959	TIGR00126				1		
arCOG01327	F	Nucleotide transport and metabolism	Pnp	Purine nucleoside phosphorylase	COG00005	pfam01048		TIGR01694				1		
arCOG00695	F	Nucleotide transport and metabolism	SsnA	Cytosine deaminase or related metal-dependent hydrolase	COG00402	pfam01979	cd01298	TIGR03314				2	2	
arCOG01566	F	Nucleotide transport and metabolism	NrnA	nanoRNase/pAp phosphatase, hydrolyzes c-di-AMP and oligoRNAs	COG00618	pfam02254,pfam01368,pfam02272							1	
arCOG00777	Q	Secondary metabolites biosynthesis, transport and catabolism	PaaI	HGG motif-containing thioesterase, possibly involved in aromatic compound	COG02050	pfam03061	cd03443	TIGR00369	6			1		
arCOG03570	Q	Secondary metabolites biosynthesis, transport and catabolism	-	SAM-dependent methyltransferase	COG00500	pfam12847	cd02440	TIGR02021	4			1		
arCOG01792	Q	Secondary metabolites biosynthesis, transport and catabolism	-	SAM-dependent methyltransferase	COG00500	pfam08241	cd02440	TIGR01934	4					
arCOG01521	Q	Secondary metabolites biosynthesis, transport and catabolism	-	Phytoene dehydrogenase or related enzyme	COG01233	pfam13450			4					
arCOG04347	Q	Secondary metabolites biosynthesis, transport and catabolism	-	SAM-dependent methyltransferase	COG00500	pfam08241	cd02440	TIGR01934	3					
arCOG03914	Q	Secondary metabolites biosynthesis, transport and catabolism	SufI	Multicopper oxidase	COG02132	pfam07732	cd11024	TIGR02376	3					
arCOG01523	Q	Secondary metabolites biosynthesis, transport and catabolism	-	Phytoene dehydrogenase or related enzyme	COG01233	pfam01593		TIGR03467	3			1	1	
arCOG01778	Q	Secondary metabolites biosynthesis, transport and catabolism	-	SAM-dependent methyltransferase	COG00500	pfam08241	cd02440	TIGR02072	3					
arCOG01781	Q	Secondary metabolites biosynthesis, transport and catabolism	-	SAM-dependent methyltransferase	COG00500	pfam08241	cd02440	TIGR02072	2					
arCOG00235	Q	Secondary metabolites biosynthesis, transport and catabolism	MhpD	2-keto-4-pentenoate hydratase/2-oxohepta-3-ene-1,7-dioic acid hydratase (	COG00179	pfam01557		TIGR02303	1			1	1	
arCOG00696	Q	Secondary metabolites biosynthesis, transport and catabolism	HutI	Imidazolonepropionase or related amidohydrolase	COG01228	pfam13147	cd01296	TIGR01224				2	1	
arCOG06106	Q	Secondary metabolites biosynthesis, transport and catabolism	-	Predicted ring-cleavage extradiol dioxygenase	COG02514	pfam12681,pfa	cd07255,cd072	TIGR03211			1			
arCOG05015	Q	Secondary metabolites biosynthesis, transport and catabolism	-	SAM-dependent methyltransferase	COG00500	pfam12847	cd02440	TIGR03534				1		
arCOG01943	Q	Secondary metabolites biosynthesis, transport and catabolism	PncA	Amidase related to nicotinamidase	COG01335	pfam00857	cd00431	TIGR03614					1	
arCOG04786	Q	Secondary metabolites biosynthesis, transport and catabolism	-	1,2-phenylacetyl-CoA epoxidase, catalytic subunit	COG03396	pfam05138		TIGR02158				1		
arCOG06169	Q	Secondary metabolites biosynthesis, transport and catabolism	-	Arylsulfotransferase family protein		pfam05935							1	
arCOG01402	Q	Secondary metabolites biosynthesis, transport and catabolism	AglP	SAM-dependent methyltransferase	COG00500	pfam05050		TIGR01444					2	
POORLY CHARACTERIZED														
arCOG04364	S	Function unknown	-	Uncharacterized protein	COG01772	pfam04407			8					
arCOG05839	S	Function unknown	-	Uncharacterized protein				TIGR04292	6					
arCOG02142	S	Function unknown	-	Pheromone shutdown protein TraB, contains GTxH motif	COG01916	pfam01963		TIGR00261	5	1		1	1	
arCOG02884	S	Function unknown	-	Membrane associated protein with extracellular Ig-like domain, a componen	COG04743				5					
arCOG10153	S	Function unknown	-	Uncharacterized membrane protein, DUF4112 family		pfam13430			5					
arCOG04596	S	Function unknown	-	Uncharacterized membrane protein					4					
arCOG03124	S	Function unknown	-	Pentapeptide repeats containing protein	COG01357	pfam00805,pfam00805			4	1		3	2	
arCOG02087	S	Function unknown	-	Predicted membrane protein	COG01470	pfam10633,pfam13620,pfam10633			4		1	2		
arCOG10338	S	Function unknown	-	Uncharacterized protein	SC.00914	pfam13517,pfam13517,pfam13517,pfam13517			4		1			
arCOG10444	S	Function unknown	-	Uncharacterized protein					4			4		
arCOG06533	S	Function unknown	-	Uncharacterized membrane protein					4					
arCOG04579	S	Function unknown	-	Uncharacterized protein, DUF2071 family	COG03361	pfam09844			4					
arCOG05495	S	Function unknown	-	Uncharacterized protein, contains N-terminal coiled-coil domain	COG04911	pfam09969			3	1				
arCOG04412	S	Function unknown	-	Uncharacterized protein	COG04004				3	1				
arCOG03232	S	Function unknown	-	VanZ like family protein	COG05652	pfam04892			3		1			
arCOG02508	S	Function unknown	-	Secreted protein, with PKD repeat domain	COG03291	pfam00801,pfa	cd00146,cd001	TIGR00864	3		2		1	
arCOG02491	S	Function unknown	-	WD40 repeats containing protein	COG02319	pfam13360			3				1	
arCOG07412	S	Function unknown	-	Uncharacterized protein					2	1				
arCOG05338	S	Function unknown	-	Uncharacterized protein					2		1	2	1	
arCOG03888	S	Function unknown	-	Uncharacterized membrane protein					2					
arCOG01119	S	Function unknown	-	GYD domain, alpha/beta barrel superfamily	COG04274	pfam08734			2					
arCOG08211	S	Function unknown	-	Uncharacterized membrane protein	SC.00448				2					
arCOG08977	S	Function unknown	-	Uncharacterized membrane protein	COG04270				2			2	1	
arCOG11014	S	Function unknown	-	HTH-domain containing transcriptional regulator	SC.00234	pfam06224			1					
arCOG03633	S	Function unknown	-	Uncharacterized membrane protein YckC, RDD family	COG01714	pfam06271			1	1				
arCOG08231	S	Function unknown	-	Periplasmic protein with immunoglobin-like fold		pfam00932			1	1			1	
arCOG05330	S	Function unknown	-	Uncharacterized protein					1	1				
arCOG02081	S	Function unknown	-	Predicted membrane protein	COG01470	pfam10633			1			1	2	
arCOG02206	S	Function unknown	-	Uncharacterized membrane protein					1		1	1		
arCOG05022	S	Function unknown	-	Uncharacterized protein, contains PQ loop repeat	COG04095				1			1	1	
arCOG06958	S	Function unknown	-	Uncharacterized protein					1			2		
arCOG06742	S	Function unknown	-	NosD-like cell surface protein					1					
arCOG08643	S	Function unknown	-	Secreted protein with C-terminal PEFG domain					1					
arCOG02527	S	Function unknown	-	Secreted protein, with PKD repeat domain	COG03291				1					
arCOG02177	S	Function unknown	-	Uncharacterized membrane protein	COG01967	pfam01889					1	2	1	
arCOG04565	S	Function unknown	-	Predicted membrane protein, DUF368 family	COG02035	pfam04018					1			
arCOG05517	S	Function unknown	-	Uncharacterized protein						1	1			
arCOG04619	S	Function unknown	-	Uncharacterized membrane protein, contains bPH2 (bacterial pleckstrin hom	COG03428	pfam03703,pfam03703,pfam03703				1		1		
arCOG10198	S	Function unknown	-	Uncharacterized membrane protein, DUF2899 family		pfam11449						1		
arCOG04076	S	Function unknown	-	Uncharacterized protein, DUF359 family	COG01909	pfam04019				1		1		
arCOG07813	S	Function unknown	-	LamG-like jellyroll fold domain	SC.00184	pfam13385							1	
arCOG02546	S	Function unknown	-	Secreted protein, contains PKD repeats and vWA domain	COG03291	pfam05048,pfa	cd00146,cd001	TIGR04247,TIGR04247,TIGR00864,TIGR04213		1		2		
arCOG03606	S	Function unknown	-	Uncharacterized protein	SC.00401					1				
arCOG01917	S	Function unknown	-	Zn-ribbon domain containing protein						1		1		
arCOG02761	S	Function unknown	-	Uncharacterized protein, DUF302 family	COG03439	pfam03625	cd14797			1				
arCOG04321	S	Function unknown	-	Uncharacterized protein DUF711 family, similar to ribonucleotide reductase	COG02848	pfam05167	cd08025			1				
arCOG05852	S	Function unknown	-	HEAT repeats containing protein	COG01413					1				
arCOG01159	S	Function unknown	-	Uncharacterized archaeal coiled-coil protein	COG01340			TIGR02168		1		4		
arCOG08673	S	Function unknown	-	Uncharacterized protein						1			1	
arCOG05351	S	Function unknown	-	Uncharacterized membrane protein		pfam11433				1				
arCOG01907	S	Function unknown	AIM24	Uncharacterized protein, AIM24 family	COG02013	pfam01987		TIGR00266				1		
arCOG04214	S	Function unknown	-	Uncharacterized protein	COG00432	pfam01894		TIGR00149					1	
arCOG01330	S	Function unknown	-	Uncharacterized membrane protein	COG02426	pfam06695					1	2		
arCOG04269	S	Function unknown	-	Protein, predicted to be involved in DNA repair	COG01602	pfam04894,pfam04895					1			
arCOG02565	S	Function unknown	-	Secreted uncharacterized protein	COG05276						1			
arCOG06429	S	Function unknown	-	Uncharacterized protein, contains NRDE domain	COG03332	pfam05742					1			
arCOG04693	S	Function unknown	-	Uncharacterized protein								1		
arCOG03949	S	Function unknown	-	Uncharacterized membrane protein	COG03815	pfam09858						2		
arCOG07680	S	Function unknown	-	Uncharacterized membrane protein	COG02832	pfam04304							1	
arCOG11882	S	Function unknown	-	Uncharacterized membrane protein								2		
arCOG04662	S	Function unknown	-	Uncharacterized membrane protein	SC.00128								1	
arCOG05345	S	Function unknown	-	Uncharacterized protein									1	
arCOG04500	S	Function unknown	-	Uncharacterized protein with Ig-like domain								1	1	
arCOG02170	S	Function unknown	-	Uncharacterized protein	SC.00300	pfam13559						1		
arCOG09426	S	Function unknown	-	Uncharacterized membrane protein								2		
arCOG03118	S	Function unknown	-	Membrane protein DegA family	COG01238	pfam09335						1	1	
arCOG02561	S	Function unknown	-	WD40 repeats containing protein	COG02319	pfam13191,pfa	cd00200,cd002	TIGR02794,TIGR02800,TIGR02800				1		
arCOG00620	S	Function unknown	-	Uncharacterized conserved DUF39 domain fused to CBS domain	COG01900	pfam01837,pfa	cd04605	TIGR03287,TIGR01302					1	
arCOG04574	S	Function unknown	LemA	Uncharacterized protein	COG01704	pfam04011							1	
arCOG04811	S	Function unknown	-	Uncharacterized membrane protein, Fun14 family	COG02383								1	
arCOG04848	S	Function unknown	-	Uncharacterized protein, DUF1722 family	COG03272	pfam04463,pfam08349							1	
arCOG06436	S	Function unknown	-	Uncharacterized protein, YdiU/UPF0061 family	COG00397	pfam02696							1	

arCOG06646	S	Function unknown	-	Uncharacterized membrane protein, YccA/Bax inhibitor family	COG04760	pfam12811							1
arCOG14289	S	Function unknown	-	Uncharacterized protein									1
arCOG10710	S	Function unknown	-	Uncharacterized protein									2
arCOG01908	S	Function unknown	AIM24	Uncharacterized protein, AIM24 family	COG02013	pfam01987						1	
arCOG03873	S	Function unknown	-	Uncharacterized membrane protein	COG02855	pfam03601		TIGR00698				1	
arCOG05340	S	Function unknown	-	Uncharacterized membrane protein								1	
arCOG05874	S	Function unknown	-	Uncharacterized protein								1	
arCOG06493	S	Function unknown	-	HEAT repeats containing protein	COG01413	pfam13646,pfam03130						1	
arCOG10568	S	Function unknown	-	Uncharacterized membrane protein								1	
arCOG02539	S	Function unknown	-	Secreted protein, with PKD repeat domain	COG03291							2	
arCOG10301	S	Function unknown	-	Uncharacterized protein	COG03011	pfam04134						2	
arCOG03682	R	General function prediction only	SPS1	Membrane associated serine/threonine protein kinase	COG00515	pfam00069	cd14014		6	4	2		1
arCOG06897	R	General function prediction only	-	Predicted ATP-grasp enzyme	COG03919	pfam15632		TIGR01369	6		1		
arCOG00347	R	General function prediction only	-	Archaeal enzyme of ATP-grasp superfamily	COG01938	pfam09754		TIGR00161	5	1			1
arCOG04444	R	General function prediction only	-	ACT domain containing protein	COG04747	pfam01842,pfa	cd04908,cd04908		5	2	1		
arCOG02293	R	General function prediction only	-	HAD superfamily hydrolase	COG00637	pfam13419	cd01427	TIGR02009	5		1		
arCOG04303	R	General function prediction only	-	Uncharacterized Rossmann fold enzyme	COG01634	pfam01973	cd07995		4	1			1
arCOG02174	R	General function prediction only	-	Predicted exporter of the RND superfamily	COG01033	pfam03176,pfam03176		TIGR00921	4			7	2
arCOG01285	R	General function prediction only	-	OB-fold domain and Zn-ribbon containing protein, possible acyl-CoA-binding	COG01545	pfam12172,pfam01796			4				
arCOG00517	R	General function prediction only	-	Rhodanese Homology Domain fused to Zn-dependent hydrolase of beta-lacta	COG00491	pfam00753		TIGR03413	3			1	2
arCOG00313	R	General function prediction only	Sco1	Cytochrome oxidase Cu insertion factor, SCO1/SenC/PrrC family	COG01999	pfam02630	cd02968		3	1			2
arCOG03096	R	General function prediction only	-	NAD dependent epimerase/dehydratase family enzyme	COG01090	pfam01370,pfa	cd05242	TIGR01777	3	1			
arCOG02812	R	General function prediction only	-	Bacteriorhodopsin	COG05524	pfam01036			3				
arCOG00352	R	General function prediction only	Nog1	GTP-binding protein, GTP1/Obg family	COG01084	pfam01926	cd01897	TIGR02729	3				
arCOG02259	R	General function prediction only	-	Zn-finger domain containing protein	SC.00304				3				
arCOG03439	R	General function prediction only	-	vWFa domain containing protein					2		1	3	2
arCOG07416	R	General function prediction only	MhpC	Alpha/beta superfamily hydrolase	COG00596	pfam12695		TIGR03695	2			1	1
arCOG02889	R	General function prediction only	-	Predicted deacylase	COG03608	pfam04952	cd06252	TIGR02994	2				
arCOG01622	R	General function prediction only	MviM	Predicted dehydrogenase	COG00673	pfam01408,pfam02894			2				1
arCOG03991	R	General function prediction only	-	PKD repeats containing protein	COG03291	pfam00801	cd00146	TIGR00864,TIGR04213	1	1		1	
arCOG01651	R	General function prediction only	-	Alpha/beta superfamily hydrolase	COG01073	pfam08538			1				
arCOG01425	R	General function prediction only	-	RecB family nuclease with coiled-coil N-terminal domain	SC.00001				1	1		1	
arCOG01141	R	General function prediction only	-	ICC-like phosphoesterase	COG00622	pfam12850	cd00841	TIGR00040	1				
arCOG01395	R	General function prediction only	-	Lipid-A-disaccharide synthase related glycosyltransferase	COG01817	pfam04007			1				
arCOG01648	R	General function prediction only	MhpC	Alpha/beta superfamily hydrolase	COG00596	pfam12697		TIGR02427	1				
arCOG02164	R	General function prediction only	-	Predicted transglutaminase-like protease	COG01800	pfam04473			1				
arCOG00614	R	General function prediction only	SpoIVFB	Zn-dependent protease	COG01994		cd06158		1				
arCOG03015	R	General function prediction only	-	Uncharacterized protein YbjT, contains NAD(P)-binding and DUF2867 domair	COG00702	pfam13460	cd05271	TIGR03466	1			1	
arCOG02625	R	General function prediction only	-	Predicted metal-dependent hydrolase	COG01451	pfam01863			1				
arCOG06095	R	General function prediction only	-	ABC-2 family transporter protein	COG01277	pfam12679			1				
arCOG01136	R	General function prediction only	ARC1	EMAP domain RNA-binding protein	COG00073	pfam01588	cd02800	TIGR00399	1			1	
arCOG01213	R	General function prediction only	Cof	HAD superfamily hydrolase	COG00561	pfam08282	cd01427	TIGR01487	1			2	1
arCOG00833	R	General function prediction only	RimI	Acetyltransferase (GNAT) family	COG00456	pfam00583	cd04301	TIGR01575		1			1
arCOG02452	R	General function prediction only	-	KaiC-like protein ATPase	COG00467					1			1
arCOG04430	R	General function prediction only	-	HD superfamily phosphohydrolase	COG01078	pfam01966	cd00077					2	1
arCOG02680	R	General function prediction only	-	Uncharacterized archaeal Zn-finger protein	COG01326					1	1		1
arCOG09205	R	General function prediction only	-	Predicted metal-dependent hydrolase	COG01547	pfam03745					1		
arCOG09415	R	General function prediction only	-	Membrane associated metal-binding domain fused to Reeler domain		pfam02014	cd08544				1	2	2
arCOG00626	R	General function prediction only	TlyC	Hemolysins or related protein containing CBS domains	COG01253	pfam01595,pfa	cd04590	TIGR03520				1	
arCOG02291	R	General function prediction only	-	HAD superfamily hydrolase	COG01011	pfam13419	cd01427	TIGR02252		1	1		
arCOG00979	R	General function prediction only	-	Predicted O-methyltransferase YrrM	COG04122	pfam13578	cd02440			1	1		
arCOG01619	R	General function prediction only	ARA1	Aldo/keto reductase, related to diketogulonate reductase	COG00656	pfam00248	cd06660	TIGR01293			1		
arCOG01641	R	General function prediction only	-	Predicted RNA-binding protein, contains TRAM domain	COG03269							1	
arCOG00504	R	General function prediction only	-	Metal-dependent hydrolase of the beta-lactamase superfamily II	COG00491	pfam00753		TIGR03413					1
arCOG00348	R	General function prediction only	-	Archaeal enzyme of ATP-grasp superfamily	COG02047	pfam09754		TIGR00162				1	
arCOG04055	R	General function prediction only	-	SHS2 domain protein implicated in nucleic acid metabolism	COG01371	pfam01951						1	1
arCOG00893	R	General function prediction only	-	Predicted metal-dependent hydrolase (urease superfamily)	COG01831	pfam01026						2	1
arCOG02579	R	General function prediction only	-	Fe-S-cluster containining protein	COG00727								1
arCOG04212	R	General function prediction only	-	Predicted DNA-binding protein with PD1-like DNA-binding motif	COG01661	pfam03479	cd11378			1			
arCOG06747	R	General function prediction only	-	SIR2 superfamily protein	COG00846	pfam02146,pfa	cd01406			1			
arCOG06769	R	General function prediction only	-	GTPase SAR1 family domain fused to Leucine-rich repeats domain	COG01100	pfam13855,pfa	cd00116,cd099	TIGR00231		1			
arCOG00499	R	General function prediction only		Metal-dependent hydrolase of the beta-lactamase superfamily	COG01235	pfam12706		TIGR02651		1		1	
arCOG01225	R	General function prediction only	-	GTPase SAR1 or related small G protein	COG01100	pfam03029	cd02027,cd00880						1
arCOG03450	R	General function prediction only	-	Predicted surface protease of transglutaminase family						1			
arCOG04624	R	General function prediction only	-	Multimeric flavodoxin WrbA	COG00431							1	1
arCOG02155	R	General function prediction only	-	Protein implicated in RNA metabolism, contains PRC-barrel domain	COG01873	pfam05239				1			
arCOG01848	R	General function prediction only	WbbJ	Acetyltransferase (isoleucine patch superfamily)	COG00110	pfam00132,pfa	cd03358	TIGR03570					1
arCOG06534	R	General function prediction only	-	Cohesin domain containing secreted protein		pfam00963	cd08547				1		
arCOG04179	R	General function prediction only	PDCD5	DNA-binding TFAR19-related protein, PDSD5 family	COG02118	pfam01984					1		1
arCOG00601	R	General function prediction only	-	Protein containing two CBS domains (some fused to C-terminal double-strand	COG00517	pfam00571,pfa	cd04633,cd046	TIGR01302,TIGR01302			1		
arCOG02289	R	General function prediction only	-	Uncharacterized protein, DUF169 family	COG02043	pfam02596					1		
arCOG03994	R	General function prediction only	-	PKD domain containing protein	COG03291	pfam05048,pfa	cd14251,cd001	TIGR04247,TIGR04275,TIGR04275,TIGR04275,TIGR04275,TIGR042			1		
arCOG03038	R	General function prediction only	-	TPR repeats containing protein	COG00457						1	1	1
arCOG05195	R	General function prediction only	-	TPR repeats containing protein	COG00457	pfam13414	cd00189	TIGR02521				2	
arCOG00353	R	General function prediction only	HflX	GTP-binding protein protease modulator	COG02262	pfam13167,pfa	cd01878	TIGR03156				1	1
arCOG01043	R	General function prediction only	-	Predicted RNA binding protein with dsRBD fold	COG01931	pfam01877						1	1
arCOG04589	R	General function prediction only	CinA	Uncharacterized protein (competence- and mitomycin-induced)	COG01546	pfam02464		TIGR00199				1	1
arCOG05137	R	General function prediction only	-	TPR repeats containing protein	COG00457	pfam13414,pfa	cd00189,cd001	TIGR02917				1	
arCOG01084	R	General function prediction only	MazG	Predicted pyrophosphatase	COG01694	pfam03819	cd11535						1
arCOG00497	R	General function prediction only	-	Zn-dependent hydrolase of the beta-lactamase fold	COG02220	pfam13483							1
arCOG00606	R	General function prediction only	-	CBS domain	COG00517	pfam00478	cd04623	TIGR01302					1
arCOG04469	R	General function prediction only	-	Tripartite tricarboxylate transporter (TTT) class transporter	COG01784	pfam01970							1
arCOG03032	R	General function prediction only	-	TPR repeats containing protein	COG00457	pfam13414,pfa	cd00189	TIGR02917				1	
arCOG10597	R	General function prediction only	-	Hemocyanin family protein, binds copper ions		pfam00264						1	
arCOG00498	R	General function prediction only	-	Metal-dependent hydrolase of the beta-lactamase superfamily II	COG00491	pfam00753		TIGR03413				1	1
arCOG00500	R	General function prediction only	ElaC	Metal-dependent hydrolase of the beta-lactamase superfamily	COG01234	pfam12706		TIGR02651				1	
arCOG02642	R	General function prediction only	PerM	Predicted PurR-regulated permease PerM	COG00628	pfam01594		TIGR02872				2	1
arCOG01030	R	General function prediction only	-	Predicted kinase related to galactokinase and mevalonate kinase	COG02605							5	
arCOG01850	R	General function prediction only	WbbJ	Acetyltransferase (isoleucine patch superfamily)	COG00110	pfam00132	cd04647	TIGR03532					1
arCOG02303	R	General function prediction only	SurE	Predicted acid phosphatase	COG00496	pfam01975,pfam14423		TIGR00087					1
arCOG02839	R	General function prediction only	-	Uncharacterized protein related to deoxyribodipyrimidine photolyase	COG03046	pfam04244,pfam03441							1
arCOG02986	R	General function prediction only	BioY	Uncharacterized protein	COG01268	pfam02632							1
arCOG03047	R	General function prediction only	-	TPR repeats containing protein	COG00457								1
arCOG03271	R	General function prediction only	YtfP	Uncharacterized protein YtfP, gamma-glutamylcyclotransferase (GGCT)/AIG2	COG02105	pfam06094	cd06661						1
arCOG05099	R	General function prediction only	YtfP	Uncharacterized protein YtfP, gamma-glutamylcyclotransferase (GGCT)/AIG2	COG02105	pfam06094	cd06661						1
arCOG08119	R	General function prediction only	PhoX	Secreted phosphatase, PhoX family	COG03211	pfam05787							1
arCOG03561	R	General function prediction only	-	Secreted protein with beta-propeller repeat domain	COG03391	pfam01436,pfa	cd14955						2
arCOG07781	R	General function prediction only	-	Cell surface protein									2
arCOG02562	R	General function prediction only	-	Beta-propeller repeat containing protein	COG03391							1	3
arCOG03169	R	General function prediction only	-	AAA+ superfamily ATPase	COG01672	pfam01637,pfam01978						1	
arCOG07790	R	General function prediction only	-	D-glucuronyl C5-epimerase C-terminal domain related protein		pfam06662						1	
arCOG02560	R	General function prediction only	-	Secreted protein with beta-propeller repeat domain	COG03391		cd05819	TIGR03866				1	
arCOG06256	R	General function prediction only	-	Predicted esterase	COG00400	pfam02230						2	
arCOG00082	EF		PucG	Serine-pyruvate aminotransferase/archaeal aspartate aminotransferase	COG00075	pfam00266	cd06451	TIGR03301	1			1	1

Supplementary Table S7. Genes found in MG-III Metabolic Pathways

	Epipelagic MG-III	Bathy1	Bathy2
Glycolysis			
hexokinase (glk)	--	--	--
phosphoglucoisomerase(pgi)	X	X	X
phosphofructokinase (pfkA)	X	--	--
aldolase (fba/dhnA)	X	X	X
triosephosphate isomerase(tpi)	X	X	--
glyceraldehyde 3-phosphate dehydrogenase (gapA)	X	X	X
3-phosphoglycerate kinase (pgk)	X	X	X
phosphoglyceromutase (pgm/yibO)	X	X	--
enolase(eno)	X	X	--
pyruvate kinase (pykA)	--	--	--
Gluconeogenesis			
phosphoenolpyruvate synthase (ppsA)	X	X	X
enolase (eno)	X	X	--
phosphoglyceromutase (pgm)	X	--	--
3-phosphoglycerate kinase (pgk)	X	X	X
glyceraldehyde 3-phosphate dehydrogenase (gapA)	X	X	X
triosephosphate isomerase(tpi)	X	X	--
aldolase (fba/dhnA)	X	X	X
fructose bisphosphatase (suhB)	X	X	X
phosphoglucoisomerase (pgi)	X	X	X
Pentose phosphate shunt and pentose biosynthesis			
glucose-6-phosphate dehydrogenase (zwf)	--	--	--
6-phosphogluconate dehydrogenase (gnd)	--	--	--
transketolase (tktA)	X	X	X
transaldolase (talA)	--	--	--
pentose-5-phosphate-3-epimerase (yhfD)	X	X	--
ribose 5-phosphate isomerase (rpiA)	X	X	X
deoxyribose-phosphate aldolase (deoC)	--	--	--
Entner–Doudoroff pathway			
glucose-6-phosphate dehydrogenase (zwf)	--	--	--
6-phosphogluconate dehydratase (edd)	--	--	--
2-keto-3-deoxy-6-phosphogluconate aldolase (eda)	--	--	--
TCA cycle			
citrate synthase (gltA)	X	X	--
aconitase(acnA)	X	X	X
isocitrate dehydrogenase (icd)	X	--	--
a-ketoglutarate dehydrogenase (sucA, sucB)	X	--	--
succinyl-CoA synthase (sucC, sucD)	X	X	--
fumarate reductase (frdA, frdB)	X	--	--
fumarase (fumA)	X	X	X
malate dehydrogenase (mdh)	X	X	--
Purine biosynthesis			
phosphosphoribosylpyrophosphate synthase (prsA)	X	X	X
amidophosphoribosyltransferase (purF)	X	X	X
GAR synthase (purD)	X	X	--
GAR transformylase(purN/purT)	X	X	X
FGAM synthase (purL)	X	X	X
AIR synthase (purM)	X	X	X
NCAIR synthase (purK)	--	--	--
NCAIR mutase (purE)	X	--	X
SAICAR synthase (purC)	X	X	--
adenylosuccinate lyase (purB)	X	X	X
AICAR transformylase (purH2)	X	X	X
IMP cyclohydrolase (purH1)	X	X	X
adenylosuccinate synthase (purA)	X	X	--
IMP dehydrogenase (guaB)	X	--	X
GMP synthase (guaA)	X	X	--
Pyrimidine biosynthesis			
carbamoylphosphate synthase(carA, carB)	X	--	X
aspartate carbamoyltransferase (pyrB)	X	X	--
dihydroorotase (pyrC/ygeZ)	--	--	--
dihydroorotate dehydrogenase(pyrD)	X	X	X
orotate phosphoribosyl-transferase (pyrE)	X	X	X
orotidine-5'-phosphate decarboxylase (pyrF)	X	--	--
UMP kinase (pyrH)	--	--	--

NDP kinase (ndk)	X	X	X
CTP synthase (pyrG)	X	X	--
Histidine biosynthesis			
phosphosphoribosylpyrophosphate synthase (prsA)	X	X	X
ATP-phosphoribosyltransferase (hisG)	--	--	--
phosphoribosyl-ATP pyrophosphatase (hisI2)	--	--	--
phosphoribosyl-AMP cyclohydrolase(hisI1)	--	--	--
58-ProFAR isomerase (hisA)	--	--	--
imidazoleglycerol phosphate synthase (hisH, hisF)	--	--	--
imidazoleglycerol phosphate dehydratase (hisB2)	--	--	--
histidinoll phosphate aminotransferase (hisC)	X	X	--
histidinol phosphatase (hisB1)	--	--	--
histidinol dehydrogenase (hisD)	--	--	--
Branched chain amino acids biosynthesis			
threonine deaminase (ilvA)	X	--	X
acetohydroxyacid synthase (ilvB, ilvN)	--	X	--
acetohydroxyacid isomeroreductase (ilvC)	--	--	--
dihydroxyacid dehydratase (ilvD)	--	--	--
2-isopropylmalate synthase (leuA)	--	--	--
isopropylmalate isomerase (leuC, leuD)	--	--	--
3-isopropyl-malate dehydrogenase (leuB)	--	--	--
glutamate transaminase (ilvE)	X	X	--
Aromatic amino acids biosynthesis			
3-deoxyheptulosonate 7-phosphate synthase (aroG/kdsA)	--	--	--
3-dehydroquinate synthase (aroB)	--	--	--
3-dehydroquinate dehydratase (aroD)	--	--	--
shikimate dehydrogenase (aroE)	--	--	--
shikimate kinase (aroK)	--	--	--
5-enolpyruvoylshikimate 3-phosphate synthase (aroA)	--	--	--
chorismate synthase (aroC)	--	--	--
chorismate mutase (pheA1)	--	--	--
prephenate dehydratase (pheA2)	--	--	--
prephenate dehydrogenase (tyrA2)	--	--	--
tyrosine aminotransferase (tyrB)	--	--	--
antranilate synthase (trpD1, trpE)	--	--	--
antranilate phosphoribosyl-transferase (trpD2)	--	--	--
phosphoribosylantranilate isomerase (trpC2)	--	--	--
indole-glycerol phosphate synthase (trpC1)	--	--	--
tryptophan synthase (trpA, trpB)	X	--	--
Threonine biosynthesis			
aspartokinase (thrA1)	--	--	--
aspartate semialdehyde dehydrogenase (asd)	--	--	--
homoserine dehydrogenase (thrA2)	--	--	--
homoserine kinase (thrB)	--	--	--
threonine synthase (thrC)	--	--	--
Methionine biosynthesis			
aspartokinase (metL1)	--	--	--
aspartate semialdehyde dehydrogenase (asd)	--	--	--
homoserine dehydrogenase (metL2)	--	--	--
homoserine transsuccinylase (metA)	--	--	--
cystathionine g-synthase (metB)	X	X	X
b-cystathionase (metC)	--	--	--
methionine synthase (metE/metH)	--	--	--
Arginine biosynthesis			
acetylglutamate synthase (argA2)	--	--	--
acetylglutamate kinase (argB)	X	--	--
acetylglutamate phosphate reductase (argC)	--	--	--
acetylornithine aminotransferase (argD)	X	--	--
acetylornithinase (argE)	X	--	--
ornithine carbamoyltransferase (argF)	X	X	--
argininosuccinate synthase (argG)	--	--	--
argininosuccinate lyase (argH)	--	--	--
NAD biosynthesis			
aspartate oxidase (nadB)	--	--	--
quinolinate synthase (nadA)	X	X	--
quinolinate phosphoribosyltransferase (nadC)	X	X	--
nicotinic acid mononucleotide adenylyltransferase (nadD)	--	--	--
deamido-NAD ammonia ligase (nadE)	X	X	X

Riboflavin biosynthesis			
GTP cyclohydrolase II (ribA)	X	X	--
pyrimidine deaminase (ribD1)	--	--	--
pyrimidine reductase (ribD2)	X	X	X
3,4-dihydroxybutanone-4-phosphate synthase (ribB)	X	X	--
6,7-dimethyl-8-ribityllumazine synthase (ribE)	X	X	--
riboflavin synthase (ribC)	X	X	--
Siroheme biosynthesis			
Glutamyl-tRNA reductase (hemA)	--	--	--
glutamate 1-semialdehyde aminotransferase (hemL)	--	--	--
probilinogen III synthase (hemB)	--	--	--
hydroxymethylbilane synthase (hemC)	--	--	--
uroporphyrinogen III synthase (hemD)	--	--	--
uroporphirinogen methyltransferase (cysG2)	--	--	--
dimethyluroporphirinogen III dehydrogenase (cysG1)	--	--	--
Cobalamin biosynthesis			
uroporphyrinogen III methylase (cysG2)	--	--	--
precorrin-2 methylase (cbiL)	--	--	--
precorrin-3B methylase (cbiH)	--	--	--
precorrin-4 methylase (cbiF)	--	--	--
precorrin-6A reductase (cbiJ)	--	--	--
precorrin 6B methylase (cbiE)	--	--	--
precorrin 6B decarboxylase (cbiT)	--	--	--
precorrin-8x isomerase (cbiC)	--	--	--
cobyrinic acid a,c-diamide synthase (cbiA)	--	--	--
cobalt insertion protein (cobN)	--	--	--
cob(I)alamin adenosyltransferase (cobA)	--	--	--
cobyric acid synthase (cbiP)	--	--	--
cobyric acid aminotransferase (cobD)	--	--	--
cobinamide synthase (cbiB)	--	--	--
nicotinate-nucleotide:dimethylbenzimidazole phosphoribosyltransferase (cobT)	--	--	--
cobalamin synthase (cobS)	--	--	--
Biotin biosynthesis			
pimeloyl-CoA synthetase (bioW)	--	--	--
7-keto-8-aminopelargonate synthetase (bioF)	X	X	--
7,8-diaminopelargonate aminotransferase (bioA)	--	--	--
dethiobiotin synthetase (bioD)	--	--	--
biotin synthetase (bioB)	--	--	--
biotin-[acetyl-CoA carboxylase] holoenzyme synthetase (birA)	X	X	X

Cells marked with a "X" means that the protein was found.