

**SOM Table 7.** List of orthologous groups (OG) in which both *Chlamydiae* and *Planctomycetes* are heavily represented, to the exclusion of most other sequenced bacteria.

Orthologous group (OG)	Score <sup>a</sup>	Orthologous group description	# <i>Chlamyd./</i> <i>Planctomyc.</i> <sup>b</sup>	# other bacteria <sup>c</sup>
NOG05013	3.00	60 kDA outer membrane protein [kustd1987, kustd1988]	<u>8/2</u>	0
COG4284	2.88	UDP-glucose pyrophosphorylase, involved in peptidoglycan, lipid A and LPS biosynthesis [kustc0526]	<u>8/2</u>	9
NOG43380	2.86	polymorphic membrane protein precursor in <i>Chlamydiae</i> [kustd1438]	<u>7/2</u>	3
NOG06624	2.83	Lysine- and Alanine-rich, basic proteins. The <i>Chlamydia muridarum</i> representative is developmentally regulated and binds DNA.	8/1	5
COG1747	2.83	Contains a hypothetical , N-terminal domain and a transcription elongation domain GreA [kustd1459]	<u>8/1</u>	5
COG3202	2.78	ATP/ADP translocase	8/1	9
COG3952	2.77	lipid A-disaccharide synthase [kuste4276]	<u>8/1</u>	10
COG3880	2.77	transmembrane protein with conserved CXXC motifs and a UVR domain, likely interacts with COG3869 [kuste3071]	<u>8/2</u>	18
COG3869	2.77	member of the ATP:guanido phosphotransferase family, likely interacts with COG3880 [kuste3072]	<u>8/2</u>	18
COG2869	2.75	Na <sup>+</sup> -translocating NADH dehydrogenase (ubiquinone) chain C [kuste3327]	<u>8/2</u>	19
COG1347	2.75	Na <sup>+</sup> -translocating NADH dehydrogenase (ubiquinone) chain D [kustc0731]	<u>8/2</u>	19
COG2871	2.73	Na <sup>+</sup> -translocating NADH-dehydrogenase (ubiquinone) chain F [kuste3329]	<u>8/2</u>	20

COG3876	2.73	hypothetical protein	8/1	13
COG2209	2.65	Na <sup>+</sup> -translocating NADH dehydrogenase (ubiquinone) chain E [kustc0732]	<u>7/2</u>	19
COG1726	2.65	Na <sup>+</sup> -translocating NADH- dehydrogenase (ubiquinone) chain A	8/1	19
COG1427	2.64	Distant homolog to the periplasmic solute binding protein from an ABC transporter [kuste3763]	<u>7/2</u>	20
COG1579	2.63	hypothetical protein [kustc0718]	<u>8/2</u>	29

---

OG assignment was made according to the STRING database (<http://string.embl.de/>)<sup>10</sup>. The genome of the planctomycete *Gemmata obscuriglobus* was also surveyed for possible OG members, by BLASTing the sequence from one OG member against the un-annotated *G. obscuriglobus* genome nucleotide data; the best BLAST hit was then queried against the STRING database, and in 13 cases the best hits were with members of the respective OG (data not shown). ORF designations in square brackets refer to the respective ORF in the *K. stuttgartiensis* genome in those OGs where it is represented.

<sup>a</sup> The score quantifies the specificity of the orthologous group (OG) for the *Chlamydiae/Planctomycetes* superphylum. It is only calculated for OGs that occur in *Planctomycetes* as well as in *Chlamydiae*. The score is the sum of two fractions: fraction of *Chlamydiae* and *Planctomycetes* species that have the orthologous group plus two times the fraction of species outside the *Chlamydiae/Planctomycetes* that do not have the group. The maximum possible score is three, the minimum is zero. Because there are many more sequenced genomes of non-*Chlamydiae/Planctomycetes* species than of members of these phyla, the “absence” part of the score is given a higher weight (multiplied by two) than the “presence” part. Only those OGs with a score above 2.625 were considered. This threshold value was selected from manual inspection of analogous lists created for other well-established phyla like the *Proteobacteria*. In these lists a value of 2.625 was well suited to select orthologous groups representing signature families for these phyla.

<sup>b</sup> Number of *Chlamydiae* and *Planctomycete* genomes (out of 8 and 2 examined, respectively) which are represented in this OG. Underlined numbers indicate the presence of *Kuenenia stuttgartiensis* in this OG.

<sup>c</sup> Number of other (non-*Chlamydiae/Planctomycete*) bacteria (out of 157 examined) also present in this OG.