Table S3. Examples of gene families distributed in the epi- and mesopelagic clades of marine Thaumarchaeota.

Family id[1]	Gene	Epipelagic SAGs (N=4)	Nmar[2] (N=1)	Mesopelagic SAGs (N=42)
OR3550	urease subunit alpha	0	0	2
OR3005	urease subunit beta	0	0	3
OR3010	urease accessory protein UreD	0	0	3
OR2375	urease accessory protein	0	0	7
OR2177	urease accessory protein	0	0	8
OR2326	urea amidohydrolase subunit gamma	0	0	8
OR2294	urease accessory protein	0	0	8
OR2270	urease accessory protein UreD	0	0	9
OR1967	urea active transporter	0	0	10
OR2098	urease subunit alpha	0	0	10
OR3671	Deoxyribodipyrimidine photolyase	1	0	0
OR3847	twin-arginine translocation pathway signal	2	0	0
OR2460	universal stress protein	3	1	0
OR3848	catalase/peroxidase HPI	2	0	0
OR2257	ammonia monooxygenase, subunit A	2	1	6
OR2022	ammonium transporter	2	1	8
OR1861	superoxide dismutase	3	1	11
OR2115	ammonia monooxygenase operon-associated hypothetical protein	2	1	9
OR2077	ammonia monooxygenase subunit B	2	1	10
OR1607	blue (type1) copper domain-containing protein	2	1	17
OR1653	multicopper oxidase type 3	2	1	17
OR1554	ammonia monooxygenase/methane monooxygenase subunit C	2	1	19
OR1042	DSBA oxidoreductase	3	1	27
OR2116	peroxiredoxin	2	1	9
OR1393	peroxiredoxin	1	1	22
OR1993	peroxiredoxin	0	1	13
OR1659	peroxiredoxin	0	1	18
OR2405	peroxiredoxin	0	0	5
OR1050	blue (type1) copper domain-containing protein	3	1	30
OR1688	blue (type1) copper domain-containing protein	1	1	18
OR1569	blue (type1) copper domain-containing protein	0	1	18

[1]These are orthologous gene families identified by the OrthoMCL software; the family id is arbitrary.

[2]Nmar: *Nitrosopumilus maritimus* scm1, the only marine Thaumarchaeota strain in pure culture.