



**Supplementary Figure 3. 23S ribosomal RNA phylogeny.** Maximum-likelihood 23S rRNA gene tree (2520 unambiguously aligned nucleotides) showing the relationship of the Thalassoarchaea (bold and red) with other MGII (MG2-GG3 in bold and green). Along with the rRNA sequences from the fosmids from the Mediterranean DCM, we have also included several rRNA sequences from GOS scaffolds where we could identify complete 23S genes. Numbers at nodes in major branches indicate bootstrap support (shown as percentages and only those >50%) by ML in MEGA 5.10. Scale bar represents the estimated number of substitutions per site. Sampling locations: MED, Mediterranean Sea, ETP-Eastern Tropical Pacific, NAEC-North American East Coast, GI-Galapagos Islands. The GOS dataset identifiers are shown next to each GOS scaffold.