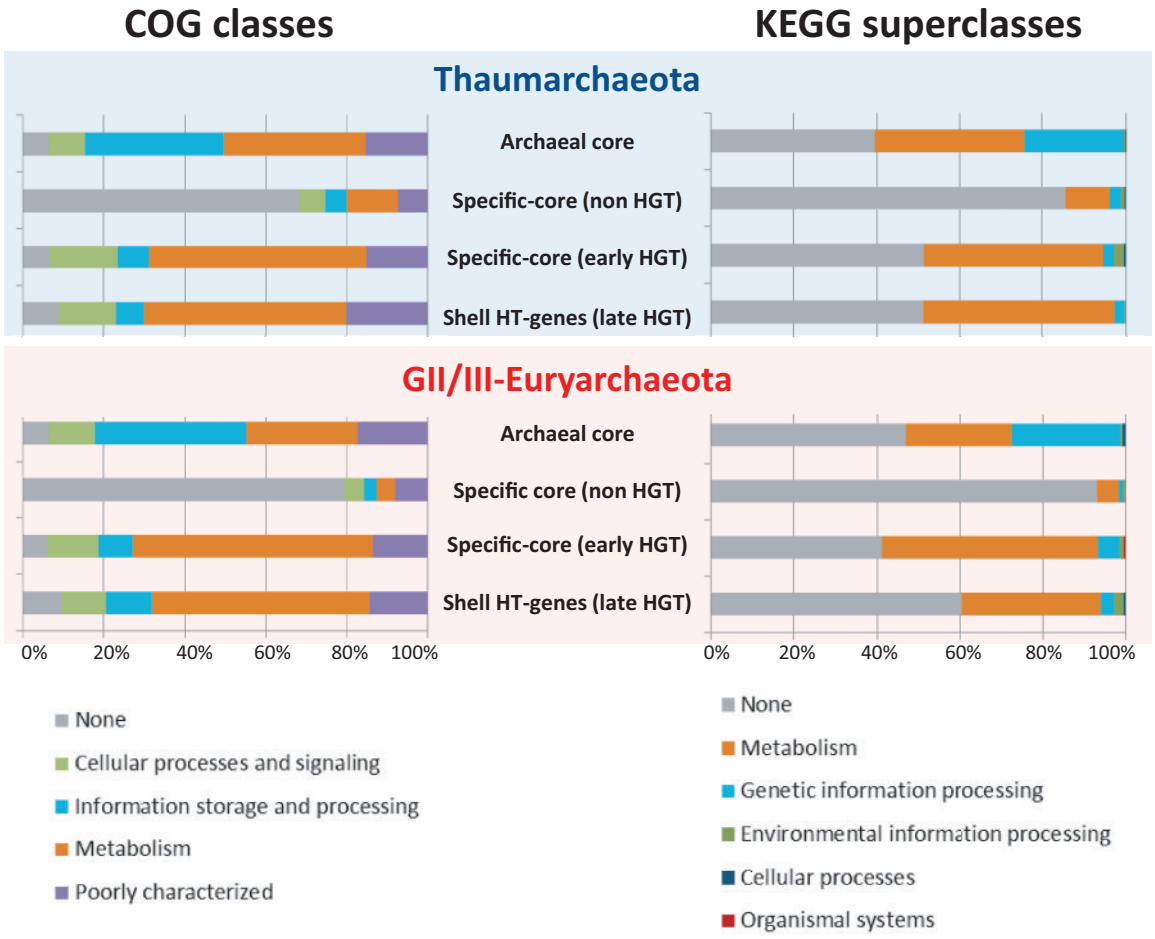

Fig. 7.—Distribution in COG classes and KEGG superclasses of deep-Mediterranean gene clusters according to their phylogenetic classification into archaeal core, lineage-specific core (non-HT-genes), early HT-genes, and accessory genes. Thaum, Thaumarchaeota; Eury, Euryarchaeota; and HTG, horizontally transferred genes.

uncultured, deep-sea Thaumarchaeota, and GII/III-Euryarchaeota. Using metagenomic fosmid libraries from deep-Mediterranean plankton, we were able to build comprehensive pangenomes for these two diverse archaeal lineages and to show, by phylogenetic analyses of all OGs, that HGT is an extensive phenomenon, with 23.9% (Thaumarchaeota) and 29.7% (GII/III-Euryarchaeota) of genes having been acquired in this way from distant donors, essentially bacteria. This level of HGT is in agreement with previous estimates based on a few fosmid and fosmid-end sequences (Brochier-Armanet et al. 2011). Even if our estimates of HGT seem high, they are indeed conservative, because we could only determine with confidence HGT cases from sufficiently resolved phylogenetic trees. Given the extent of HGT in conserved genes, it seems reasonable to hypothesize that an unknown fraction of less-conserved genes and/or genes for which sampling was too poor, which were dismissed in our analysis, might have also been acquired by HGT from distant donors. This "long-distance" HGT phenomenon is ongoing and does not affect only shell genes (recent HT-genes) but also lineage-specific core genes. This implies that a significant fraction of genes were acquired by the ancestors of marine Thaumarchaeota and GII/III-Euryarchaeota, respectively, and that a significant fraction of these transfers was also vertically inherited by the soil Thaumarchaeota branching off the marine clade (*Nitrososphaera* sharing a large fraction of those HT-genes). Although the fraction of early HT-genes was not apparently as high in GII/III-Euryarchaeota (ca. 9%), it corresponded to a prominent number of genes (416 genes; table 1) and might simply reflect the higher number of genes defined as shell. Indeed, the definition of the GII/III-Euryarchaeota pangenome was based in shared genes with a genome scaffold reconstructed from bulk short metagenome sequences (Iverson et al. 2012), which might favor the elimination of accessory genes not shared by all strains. In fact, some genes identified as late HT-genes could change to the early HT-genes as more GII/III-euryarchaeal genomes become available.