diverse groups, for the purpose of this study we considered deep-Mediterranean fosmid-derived pangenomes as representative of the two archaeal lineages. In the case of the Thaumarchaeota, we decided to include the deep-branching marine lineage to test whether some gene transfers were shared by all the marine Thaumarchaeota identified so far.

Genes for deep-Mediterranean Thaumarchaeota and GII/III-Euryarchaeota lineages were annotated and classified according to their predicted function in COG categories and KEGG classes (supplementary figs. S3 and S4, Supplementary Material online). In the case of Thaumarchaeota, genes encoding ammonium monooxygenase subunits and ammonium transporters were found to be present in equivalent numbers to single-copy genes in our Mediterranean fosmids (supplementary fig. S5, Supplementary Material online). Likewise, urease and urea transport genes were found in similar proportions. This strongly suggests that deep-Mediterranean Thaumarchaeota are ammonia oxidizers and that, similarly to their deep Arctic relatives, they utilize urea to fuel nitrification (Alonso-Saez et al. 2012). Urea degradation seems to be a metabolic feature of deep-sea Thaumarchaeota thriving in highly oligotrophic conditions, irrespective of the geographic region or local temperature, because deep-Mediterranean waters are relatively warm (14 °C on average) (Martin-Cuadrado et al. 2007). Also, all the genes that have been proposed to take part in the 3-hydroxypropionate/4-hydroxybutyrate cycle for autotrophic carbon fixation in *N. maritimus* (Walker et al. 2010) were present in the thaumarchaeal pangenome, reinforcing the idea that deep-sea planktonic Thaumarchaeota have the potential for chemolithoautotrophic growth. In contrast, despite a minimum of nine complete genomes were represented in the GII/III-Euryarchaeota data set, genes encoding proteorhodopsin homologs were not detected. This absence suggests that these GII-Euryarchaeota are genuine deep-sea dwellers that differ from their surface, proteorhodopsin-containing, counterparts (Frigaard et al. 2006; Iverson et al. 2012). They are most likely heterotrophic given the abundance of genes involved in amino acid, carbohydrate, and lipid transport and metabolism (see below). This is in agreement with deep-sea metatranscriptomic studies showing high levels of GII-Euryarchaeota amino acid transporter transcripts (Baker et al. 2013).

Determining Categories of Core, Lineage-Specific Core, and Shell Genes in Archaeal Pangenomes

Using fosmid sequences in metagenomic studies offers the advantage (shared with single-cell genomes when they are not too partial) of having access to sets of genes that are physically linked in a genome, therefore allowing the identification of accessory genes that are rare or present only in a subset of strains and that might be overlooked when reconstructing composite scaffolds from bulk short metagenomic sequences (Iverson et al. 2012). Starting from our fosmid sequences, we could thus define collections of OG clusters representing deep-Mediterranean Thaumarchaeota and GII/III-Euryarchaeota pangenomes. Subsequently, we classified them into core archaeal genes (universal genes and genes shared by all archaea), lineage-specific core genes (genes shared by, respectively, all—or all but one in the case of Thaumarchaeota, to accommodate single-lineage losses—archaeal genomes), and shell or accessory genes (only present in one or a reduced subset of genomes within each lineage) (see Materials and Methods). Excluding predicted genes with no homologs (orphans), a total of 2,098 and 3,527 OG clusters were identified, respectively, for the thaumarchaeal and GII/III-euryarchaeal pangenomes (table 1). Some of them were universal genes or genes shared by all archaea (629 and 552 for, respectively, thaumarchaeal and GII/III-euryarchaeal pangenomes). To define Thaumarchaeota-specific core genes, we considered OGs shared by our deep-Mediterranean fosmids and the genomes of their closest phylogenetic relatives from aquatic environments, namely, *N. maritimus* SCM1 (NC_010085), *C. symbiosum* A (NC_014820), and *N. limnia* SFB1 (Blainey et al. 2011) (fig. 1) to the exclusion of other archaea, which resulted in a total of 706 Thaumarchaeota-specific core genes. Likewise, for GII/III-Euryarchaeota, we used the composite genome built from surface seawater metagenome (CM001443.1) (Iverson et al. 2012), resulting in a remarkably similar number, 704, of GII/III-Euryarchaeota-specific core genes (table 1 and fig. 3). The remaining OGs present in only a subset of Thaumarchaeotal or GII/III-Euryarchaeota fosmids, having only one to three hits in the database or lacking homologs in archaea but not in other life domains were classified as shell genes. The total number of shell genes in GII/III-Euryarchaeota (2,271) was much larger than that of Thaumarchaeota (763).

Phylogenetic Identification of Early and Late HGT Events

We incorporated to our deep-Mediterranean data set representative genomic sequences covering a comprehensive taxonomic sampling of archaea (including genomes, metagenomes, and environmental fosmids), bacteria, and eukaryotes and carried out phylogenetic analyses of all OGs. These were used to refine the definition and identification of archaeal core, lineage-specific core, and shell gene classes. Whenever the query OGs were robustly nested within bacteria, eukaryotes, or other distant archaeal phyla (see criteria to define HGT events in Materials and Methods), they were considered horizontally transferred genes (HT-genes).

We identified a high HGT level in the two archaeal-lineage pangenomes, amounting to 23.9% in Thaumarchaeota and 29.7% in GII/III-Euryarchaeota (table 1 and fig. 3). These HT-genes were found in lineage-specific core and shell gene classes. HT-genes in the shell fraction nested within distant donor lineages in phylogenetic trees but were absent from complete thaumarchaeal genomes or the composite marine surface