

Supporting Information

Adam et al. 10.1073/pnas.1716667115

PNAS

Clade A
(archaeal-type CdHA)

Others
(unresolved branches of Clades E,F)

Clade E
(CdHA Firmicutes 2)

Clade E

Clade B

Clade F
(incl. CdHA Firmicutes 1)

Clade C

Clade D

Tree scale: 0.1

Fig. 51. Maximum likelihood phylogeny of all CdHA/CooS homologs (420-aa positions) calculated in IQTree under the TEST option. Values at nodes refer to bootstrap supports (100 replicates). Branches with support below 80 were collapsed. The scale bar represents the average number of substitutions per site. Sequences are color coded as follows: black (bacterial CooS), blue (bacterial CdHA), red (archaeal CooS), and green (archaeal CdHA). In the absence of experimental evidence, we tentatively classified these CdHA/CooS homologs using the following criteria: likely a CdHA if it is part of a CODH/ACS cluster or—in the case of Archaea—the only CdHA/CooS homolog in addition to a partial cluster, and likely a CooS if it is not part of a CODH/ACS cluster and/or the only homolog in the genome and displays an inconsistent phylogenetic placement. Clade names follow the conventions from ref. 6. For discussion and details on analyses, see main text and Methods section.

AAB80857_1_l-acetyl-CoA_decarboxylase_synthase_subunit_beta_cdhC_224325 [Archaeoglobus_fulgidus DSM_4304]
 AAB89186_1_l-acetyl-CoA_decarboxylase_synthase_beta_subunit_269797 [Methanococcoides_jannaschii_barkeri_str_-fusaro]
 KP765073_1_l-acetyl-CoA_decarboxylase_synthase_complex_subunit_beta_2_1700835 [Candidatus_Bathymonasetaeon_DG_70-1]
 KY35830_1_l-acetyl-CoA_decarboxylase_synthase_complex_subunit_beta_1803814 [Thermoplasmatales_methanococcus_ATCC_39073]
 ABC1916_1_l-acetyl-CoA_synthase_catalytic_subunit_beta_subunit_acetyl-CoA-decarboxylase_synthase_alpha_subunit_264732 [Morella_thermoacetica_ATCC_39073]
 AF447862_1_l-acetyl-CoA_synthase_alpha_subunit_acetyl-CoA-decarboxylase_synthase_beta_subunit_1803814 [Thermoplasmatales_methanococcus_ATCC_39073]
 BA59033_1_l-co_dehydrogenase_acetyl-CoA_synthase_complex_beta_subunit_1446466 [Candidatus_Acetothermus_aerotrophicum]
 CAJ72290_1_similar_to_CO_dehydrogenase_acetyl-CoA_synthase_alpha_subunit_1230344 [Candidatus_Xuenenia_enrichment_culture_clone_Chi]

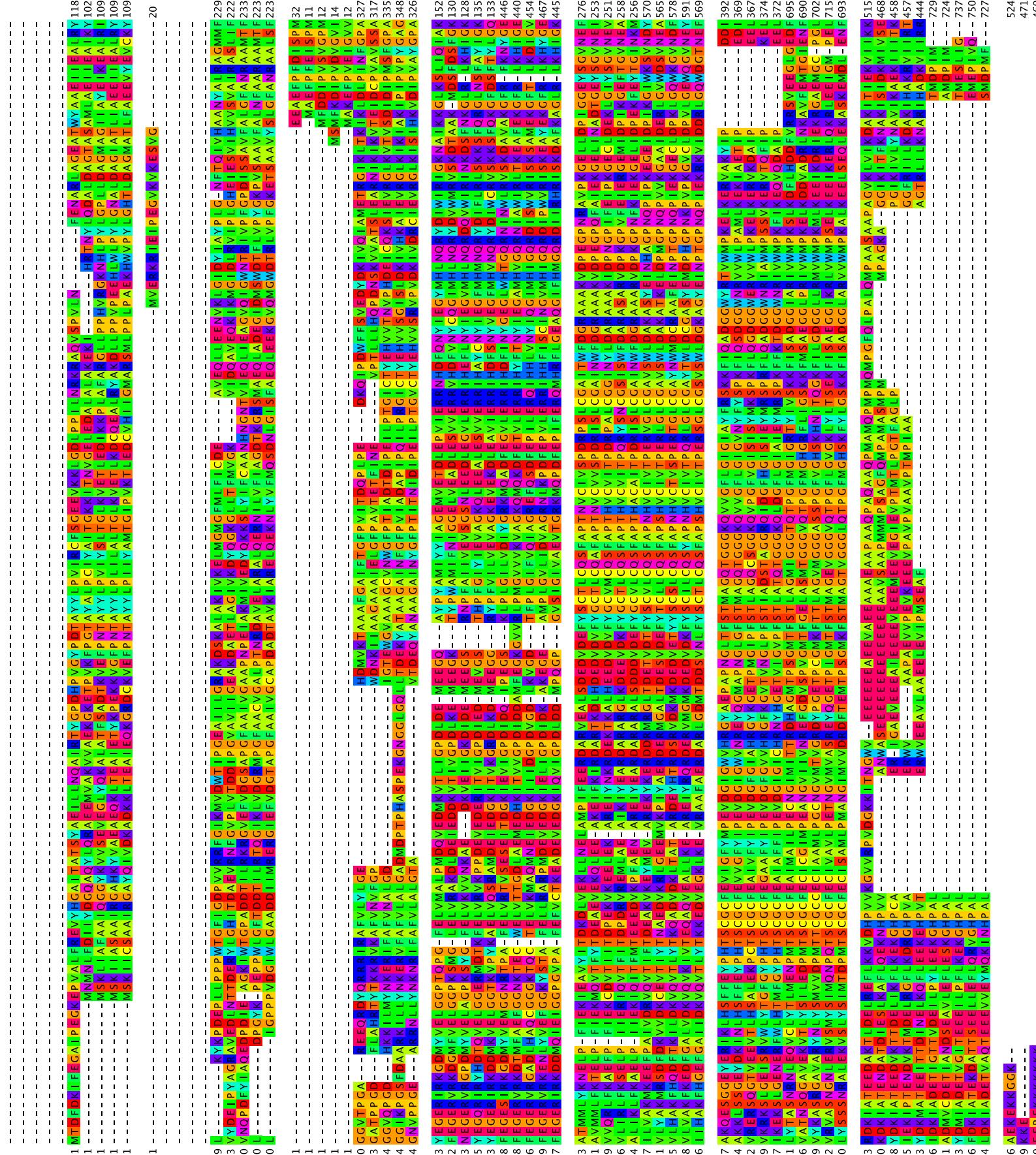


Fig. S2. Alignment of CdhC, indicating the differences between archaeal-type and bacterial-type sequences. Top five sequences represent Archaea; bottom five sequences represent Bacteria.

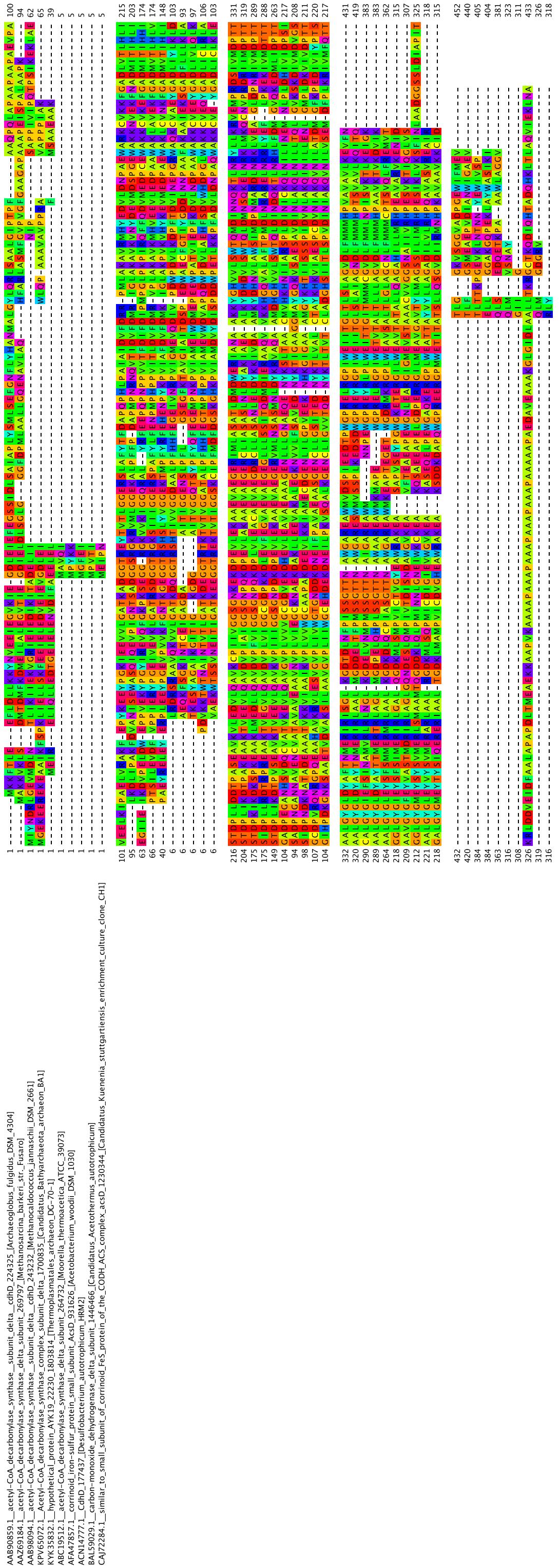


Fig. S3. Alignment of CdhD, indicating the differences between archaeal-type and bacterial-type sequences. Top five sequences represent Archaea; bottom five sequences represent Bacteria.

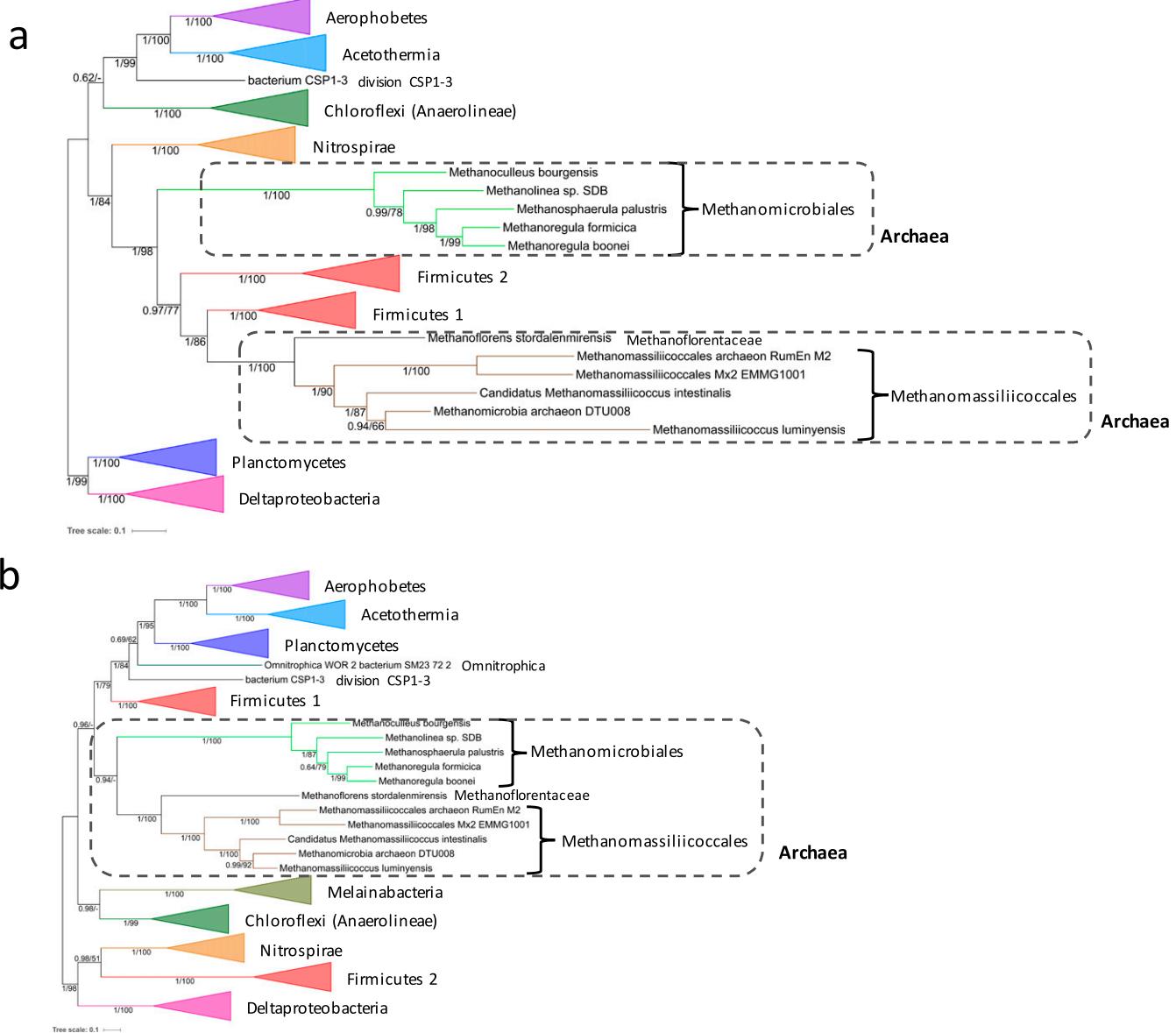


Fig. S4. Interdomain transfers of CODH/ACS components from Bacteria to Archaea. Bayesian phylogenies based on CdhC (643-aa positions) (A) and on a concatenation of the three methyltransferase module subunits (AcSE-CdhDE, 947-aa positions) (B). Values at nodes refer to posterior probabilities and bootstrap supports calculated by maximum likelihood in IQTree under the TEST option (100 replicates). Scale bars represent the average number of substitutions per site. For discussion and details on analyses, see main text and Methods.

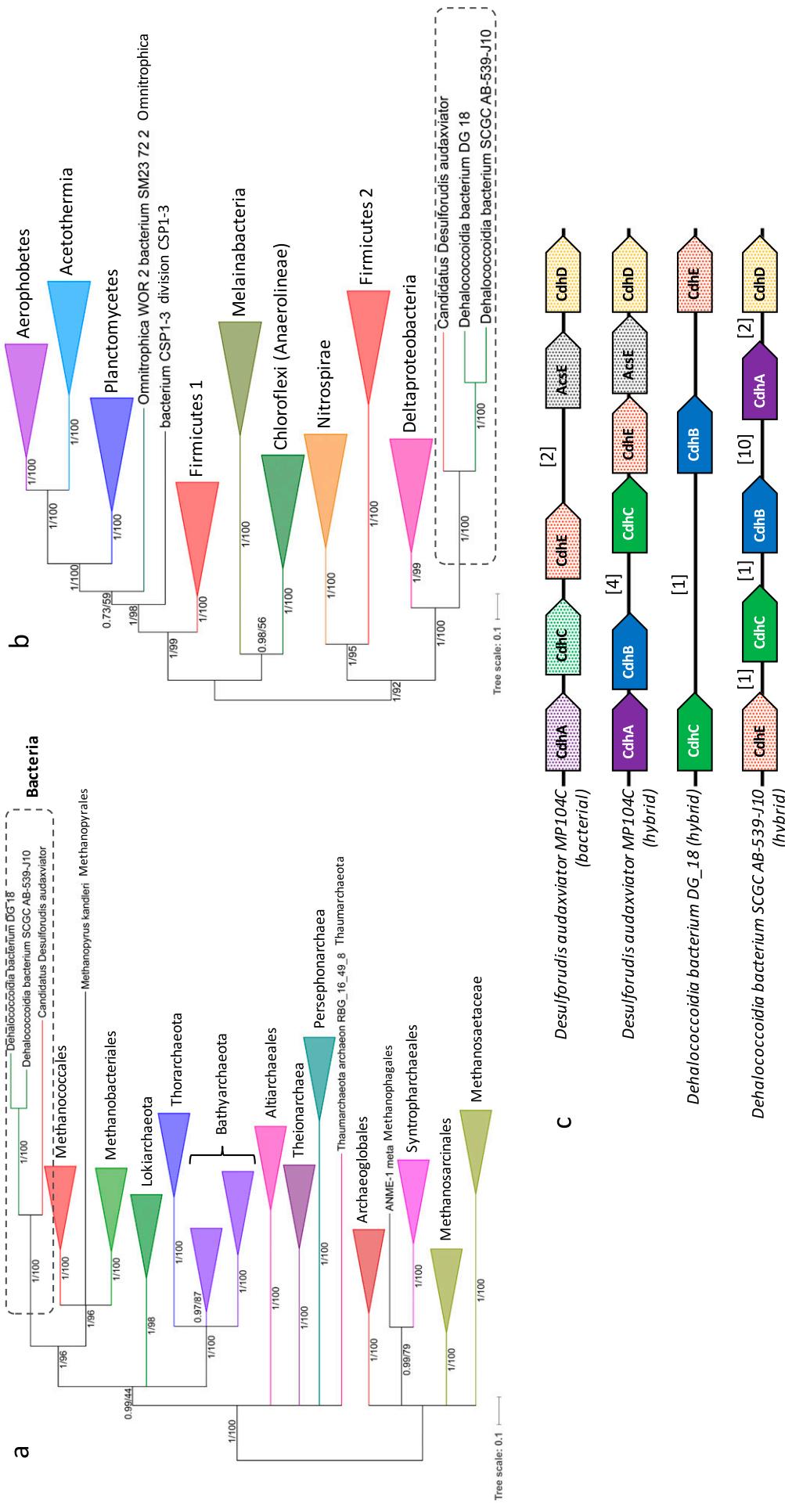


Fig. S5. (A) Interdomain transfers of CODH/ACS components from Archaea to Bacteria. Bayesian phylogeny calculated in PhyloBayes, with the CAT+GTR+I4 model based on a concatenation of the three oxidoreductase module genes (CdhABC, 1,280-aa positions). (B) Bayesian phylogeny of the three methyltransferase module subunits (AcSE–CdhDE, 953-aa positions) showing the detail of the transfer from Delta proteobacteria to *D. audaxviator* and *Dehalococcoidia*. (C) Schematic representation of the structure of the hybrid CODH/ACS clusters in *D. audaxviator* and *Dehalococcoidia*, and for comparison of the normal bacterial-type cluster in *D. audaxviator* and *Dehalococcoidia*, and for comparison of the normal bacterial-type CODH/ACS, and shaded genes are bacterial-type CODH/ACS. Numbers in square brackets refer to interspersed genes, many of which correspond to putative NADH dehydrogenases, and ferredoxins. Values at nodes refer to posterior probabilities and bootstrap supports were calculated by maximum likelihood in IQTree under the TEST option (100 replicates). Scale bars represent the average number of substitutions per site. For discussion and details on analyses, see main text and Methods.

Dataset S1. Full list of accession numbers for archaeal CODH/ACS homologs used in this analysis

Dataset S1

Some taxIDs do not correspond to public databases because the corresponding organism had not been deposited in NCBI. The same applies for some accession numbers for which a protein dataset was not available.

Dataset S2. Full list of accession numbers for bacterial CODH/ACS genes used in this analysis

Dataset S2

Some taxIDs do not correspond to public databases because the corresponding organism had not been deposited in NCBI. The same applies for some accession numbers for which a protein dataset was not available.