Check for updates

# Functional Insights of Salinity Stress-Related Pathways in Metagenome-Resolved *Methanothrix* Genomes

Maria Cristina Gagliano,[a,b] Pranav Sampara,[c] Caroline M. Plugge,[a,b] Hardy Temmink,[a,d] Dainis Sudmalis,[d] Ryan M. Ziels[c]

[a]Wetsus – European Centre of Excellence for Sustainable Water Technology, Leeuwarden, the Netherlands
[b]Laboratory of Microbiology, Wageningen University and Research, Wageningen, the Netherlands
[c]Civil Engineering, University of British Columbia, Vancouver, British Columbia, Canada
[d]Department of Environmental Technology, Wageningen University and Research, Wageningen, the Netherlands

Maria Cristina Gagliano and Pranav Sampara contributed equally to this article. Author order was determined on the basis of the study conception and the final shaping of the manuscript.

**ABSTRACT** Recently, methanogenic archaea belonging to the genus *Methanothrix* were reported to have a fundamental role in maintaining stable ecosystem functioning in anaerobic bioreactors under different configurations/conditions. In this study, we reconstructed three *Methanothrix* metagenome-assembled genomes (MAGs) from granular sludge collected from saline upflow anaerobic sludge blanket (UASB) reactors, where *Methanothrix harundinacea* was previously implicated with the formation of compact and stable granules under elevated salinity levels (up to 20 g/L Na$^+$). Genome annotation and pathway analysis of the *Methanothrix* MAGs revealed a genetic repertoire supporting their growth under high salinity. Specifically, the most dominant *Methanothrix* (MAG_279), classified as a subspecies of *Methanothrix_A harundinacea_D*, had the potential to augment its salinity resistance through the production of different glycoconjugates via the N-glycosylation process, and via the production of compatible solutes as N$^{\varepsilon}$-acetyl-$\beta$-lysine and ectoine. The stabilization and reinforcement of the cell membrane via the production of isoprenoids was identified as an additional stress-related pathway in this microorganism. The improved understanding of the salinity stress-related mechanisms of *M. harundinacea* highlights its ecological niche in extreme conditions, opening new perspectives for high-efficiency methanisation of organic waste at high salinities, as well as the possible persistence of this methanogen in highly-saline natural anaerobic environments.

**IMPORTANCE** Using genome-centric metagenomics, we discovered a new *Methanothrix harundinacea* subspecies that appears to be a halotolerant acetoclastic methanogen with the flexibility for adaptation in the anaerobic digestion process both at low (5 g/L Na$^+$) and high salinity conditions (20 g/L Na$^+$). Annotation of the recovered *M. harundinacea* genome revealed salinity stress-related functions, including the modification of EPS glycoconjugates and the production of compatible solutes. This is the first study reporting these genomic features within a *Methanothrix* sp., a milestone further supporting previous studies that identified *M. harundinacea* as a key-driver in anaerobic granulation under high salinity stress.

**KEYWORDS** anaerobic digestion, EPS, granular sludge, high salinity, *Methanothrix*, N-glycosilation, isoprenoids, methanogens, osmolytes

*M*ethanothrix (previously named *Methanosaeta*) is a methanogenic archaeal genus of which the associated members are specialists that utilize only acetate, the precursor for more than half of the methane produced on earth (1). In previous studies, *Methanothrix* species were considered to be sensitive to several inhibiting conditions, such as high concentrations of organic acids, ammonia, hydrogen, among others. (2–4).

However, more recently, multiple studies have reported on the central role of *Methanothrix* spp. in maintaining stable ecosystem functioning within perturbed anaerobic bioreactors. For instance, *Methanothrix* spp. were resistant to drastic acetate increases (5, 6), high organic acid and ammonia concentrations (7), and elevated concentrations of the long-chain fatty acid palmitate (up to 4 mmol/L) (2). *Methanothrix* was also found to be the main acetate-degrading genus within stable anaerobic reactors with efficient methane production (6, 8).

With their low excess sludge production, granular sludge systems like upflow anaerobic sludge blanket (UASB) reactors are now a widely adopted option to treat industrial wastewaters (9). *Methanothrix* spp. can constitute the initial nuclei for the formation of such granules (10–13), likely due to their ability to form filaments (13–15) as well as to the specific glycoproteic, hydrophobic sheet surrounding these filaments which stimulates bio-aggregation (16–18). In recent years, there has been a growing interest in the applicability of high-rate anaerobic reactors for treatment of saline wastewaters (19–21), amounts of which are expected to increase globally (22, 23). UASB is a promising biotechnology platform to treat saline industrial wastewater streams, for example before a desalination step for water reclamation or harvesting (24, 25). However, high $Na^+$ concentrations can cause cell lysis, inhibit methanogens and disrupt the structure of UASB granules (26, 27). Despite this, *Methanothrix harundinacea* was recently detected in compact and large-sized granules in UASB reactors at elevated salinity levels (from 5 to 20 g/L $Na^+$) (19, 28, 29). Pure cultures of *M. harundinacea* have, however, not been described to be tolerant to $Na^+$ (30).

*In-silico* analysis of the complete genome of *M. harundinacea* strain 6Ac (31) revealed several genetic loci that potentially encode for salinity-stress related pathways, such as the synthesis of compatible solutes and the production of different surface glycoconjugates constituting the extracellular polymeric substance (EPS) (32). The accumulation of compatible solutes is a well-known osmoprotection strategy by some methanogenic archaea, which mostly synthesize derivatives of $\beta$-amino acids, such as $\beta$-glutamine and $N^\varepsilon$-acetyl-$\beta$-lysine (33, 34). The secretion of EPS is also recognized as a fundamental microbial adaptation to salinity, providing osmotic tolerance and limiting dehydration (35). The composition of excreted EPS can change in response to salinity fluctuations (36–38), as has been observed in the halophilic archaeon *Haloferax volcanii* when salinity decreases (39). A similar adaptation strategy observed within halophilic archaea is the production and inclusion of isoprenoid derivatives and carotenoids into lipid membranes, which helps in maintaining their fluidity in response to changes in the osmotic conditions (40). Yet, changing the outer layer structure via the synthesis of different EPS glycoconjugates or isoprenoids are stress responses that have never been reported for methanogenic archaea. Therefore, mechanisms of adaptation to saline wastewater streams in UASB reactors remain unclear, particularly for keystone *Methanothrix* spp. (28, 32).

In our previous work, we demonstrated microbial granulation from dispersed biomass in stable UASB reactors at low salinity (LS) and high salinity (HS) (working at 5 and 20 g/L $Na^+$, respectively), with a clear dominance of *M. harundinacea* clusters in the formed granules (28, 29). Using fluorescence *in-situ* hybridization (FISH), lectin staining, and clonal sequencing analysis, we also observed different subtypes of *M. harundinacea* with distinct aggregation behavior and an EPS-glycoconjugate pattern shift in response to changes in salinity (32). Additionally, after exposing granules grown at 20 g $Na^+$/L to a salinity shock, we identified $N^\varepsilon$-acetyl-$\beta$-lysine as one of the excreted compounds, indicating the presence of anabolic pathways within the bioreactor for its formation (41). Therefore, we hypothesized that the *Methanothrix* genomes within those saline UASB reactors would contain the genetic repertoire to enable the adaptation toward high salinity. In this study, we investigated this hypothesis by analyzing samples taken in a time-series from the same saline UASB reactors using genome-resolved metagenomics. The reconstructed metagenome-assembled genomes (MAGs) were annotated for different salinity-stress related functions, and were compared in a pangenomic analysis with other representative *Methanothrix* related genomes, salinity adapted methanogens, and salinity adapted archaea. The different genomic features

identified in this work suggest that a newly discovered *M. harundinacea* subspecies can be considered a halotolerant methanogen, adding support to our previous findings identifying *Methanothrix* as a key-driver in anaerobic granulation under salinity stress.

## RESULTS AND DISCUSSION

**Classification of methanogenic metagenome-assembled genomes.** Details on the operation and performance of the LS and HS UASB reactors, working at 5 and 20 g/L Na$^+$, respectively, are presented by Sudmalis, Gagliano et al. (29), while a preliminary assessment of the microbial community based on 16S rRNA amplicon sequencing and FISH was outlined by Gagliano et al. (28). A summary of the operational and performance parameters of the two reactors is provided in Table S1. In this study, we further analyzed the archaeal community of both UASB reactors by recovering a dereplicated set of MAGs, with a focus on obtaining a higher-resolution understanding of *Methanothrix* population dynamics. Granule samples were collected from both reactors at 10 time points, starting from day 79 (after the first granules had formed in the HS reactor) until the end of the operational period (day 217; Fig. S1). The genome-resolved-metagenomics effort recovered one *Methanothrix* and two *Methanothrix*_A MAGs, based on classification with the Genome Taxonomy Database (GTDB) (42). The two *Methanothrix*_A genomes, MAG_279 and MAG_280 had a completion of 89% and 93%, respectively, and each had a redundancy of 0.7%. The single *Methanothrix* genome, MAG_281, had a completion of 94% and redundancy of 2%. *Methanothrix*_A MAG_279 shared a 99.7% average nucleotide identity (ANI) with *Methanothrix harundinacea* 56_747 (accession number LGHB00000000.1) (Fig. S2), likely belonging to the same species (43). *Methanothrix*_A MAG_280 had a maximum ANI of 88.9% with *M. harundinacea* isolate UBA475 (accession number DAXT00000000.1) across 60 *Methanothrix*_A and *Methanothrix* genomes available from the NCBI database (Fig. S2), suggesting that it may be a novel species based on an ANI cutoff of 95% (43). Finally, *Methanothrix* MAG_281 shared an ANI of 99.2% with *M. soehngenii* AS27yjCOA_157 (accession number JAAYUN000000000) (Fig. S2). The three MAGs encode the complete pathway for acetoclastic methanogenesis (Fig. S3), in agreement with the known ecological niche of *Methanothrix* (1). The recovered *Methanothrix* MAGs were compared in a pangenome analysis based on amino acid sequence similarity of genes, along with two representative *Methanothrix harundinacea*, two salinity adapted methanogens (*Methanosarcina mazei* Go1 and *Methanococcus maripaludis* C5) (44, 45) and the halophilic archaeon *Haloferax volcanii* Ds2 (46). The pangenome analysis identified 12,306 gene clusters (Fig. 1), from which three "core groups" were distinguished based on genes being systematically detected and clustered together within *Methanothrix* taxonomic groups: (i) the *Methanothrix*_A core included 610 gene clusters; (ii) the *Methanothrix* core contained 1,030 gene clusters; and (iii) the shared core by *Methanothrix*_A and *Methanothrix* constituted 418 gene clusters (Fig. 1). The *Methanothrix*_A core gene group contained a higher number of gene families connected to salinity stress in comparison to other *Methanothrix* lineages (Fig. 1). Gene clusters encoding for salinity stress functions were also found in the accessory genomes of *Methanothrix* and *Methanothrix*_A, but they were grouped separately from those of *M. mazei*, *M. maripaludis*, and *H. volcanii*, indicating a potentially unique response of the *Methanothrix* ssp. identified in this study (Supplementary Data File 1). Salinity stress-related related functions were further investigated within the reconstructed *Methanothrix* MAGs in relation to the community dynamics observed during the reactor operation, as discussed in the following sections.

**Dynamics of Methanothrix spp. in low and high salinity anaerobic reactors.** While the two *Methanothrix*_A and one *Methanothrix* MAGs were detected within both the LS and HS reactors, their abundance differed between the two systems (Fig. 2). *M. harundinacea* MAG_279 was the dominant archaeon under both salinity conditions, with an average relative abundance 12% in LS and 33% in HS, out of the total microbial population (Fig. 2). *M. harundinacea* MAG_279 remained the dominant *Methanothrix* sp. within all 10 time points analyzed for the HS reactor, while it had a similar relative
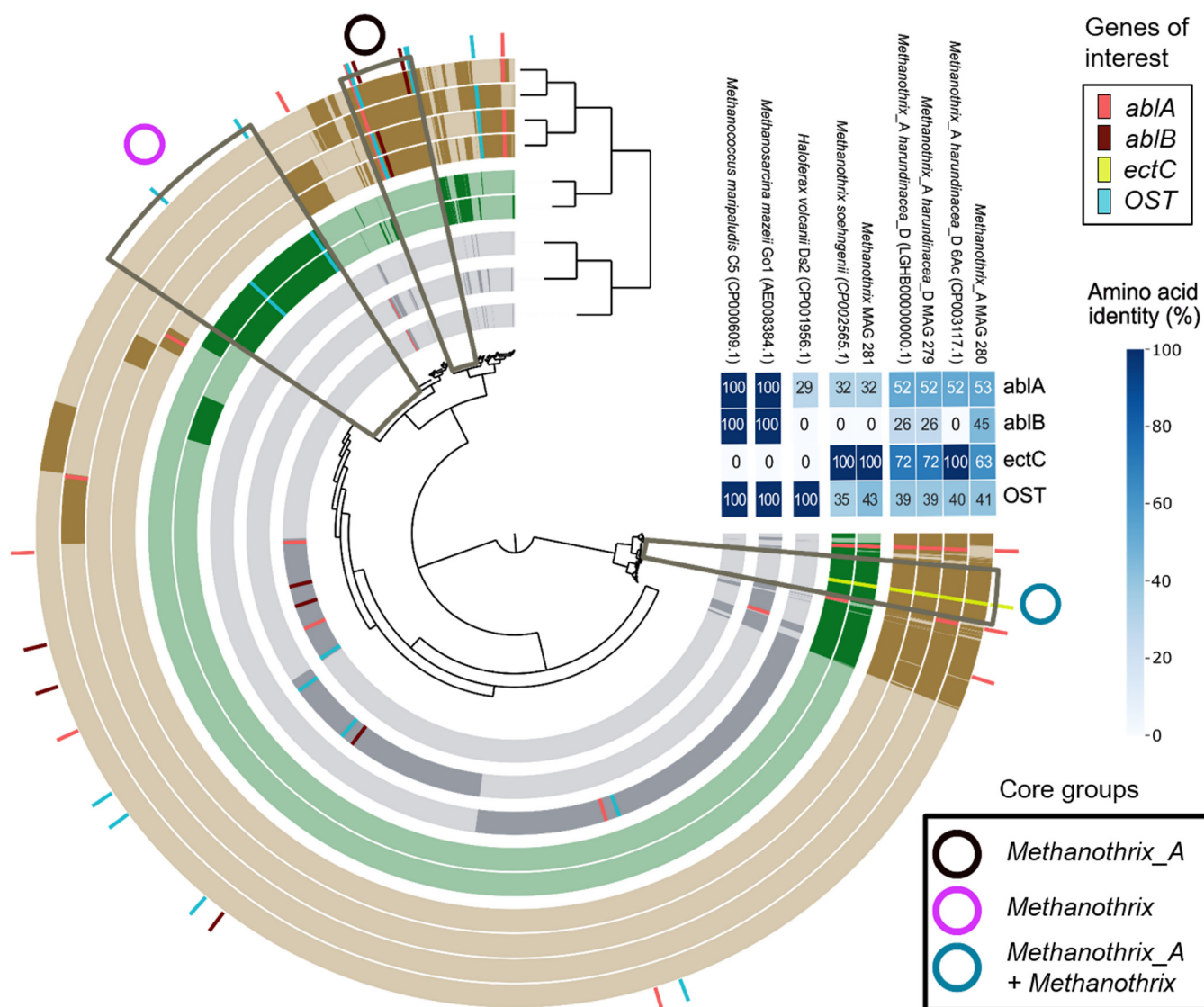
**FIG 1** Pangenome analysis of *Methanothrix* related genomes, salinity adapted methanogens and archaea from NCBI database (accession numbers in brackets), along with *Methanothrix_A harundinacea_D MAG_279*, *Methanothrix_A MAG_280*, and *Methanothrix sohegenii MAG_281*. The dendrogram on the top represents a hierarchical clustering of the genomes based on presence/absence of genes. The clustering of the genes is based on amino acid sequence similarity, as performed on Anvi'o (v7) using BLASTP. Per each genome analyzed, the core gene groups are identified with darker spots along the pangenome. Within these core groups, key genes related to salinity stress functions (as osmolytes production and EPS excretion) are highlighted (color legend on the top-right). The associated heatmap on the right shows amino acid sequence identity for those genes estimated using BLASTP. The genes are: ablA (L-lysine 2,3-aminomutase), ablB (β-lysine N(6)-acetyltransferase), ectC (L-ectoine synthase) and OST (oligosaccaryl transferases).

abundance as *Methanothrix_A MAG_280* in the LS reactor from around day 173 until the end of the experimental period (between 11% and 13%) (Fig. 2). The relative abundance of *Methanothrix_A MAG_280* in the HS reactor was substantially lower and constant between 2% to 3% (Fig. 2). The abundance of *M. soehngenii MAG_281* was negligible within both reactors ($< 1\%$), consistent with the lack of literature reporting salinity-adapted strains of this species. These results confirm that *M. harundinacea MAG_279* out-competed *Methanothrix_A MAG_280* and *M. soehngenii MAG_281* for acetate under highly saline conditions up to 20 g/L of Na$^+$. Furthermore, the ANI analysis revealed that *M. harundinacea MAG_279* shared the highest ANI with *Methanosaeta harundinacea 56_747* (GCA_001508615.1) (Fig. S2), which was identified in sediment samples collected within an oil field subsurface reservoir (47). The start-up inoculum from which *Methanothrix_A MAG_279* originated was from a full-scale UASB reactor treating industrial wastewater rich in salt ($\approx 8$ g/L Na$^+$) produced by the Shell plant in
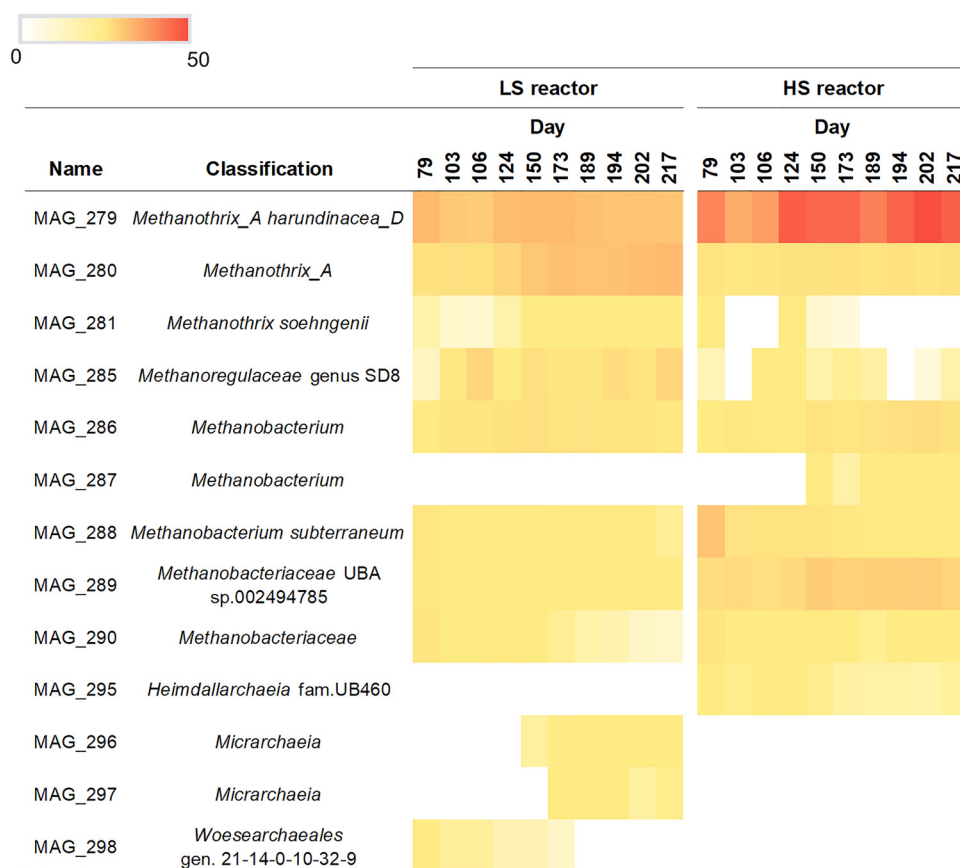
**FIG 2** Heatmap showing the relative abundance of de-replicated archaeal MAG in the LS (5 g/L Na$^+$) and HS (20 g/L Na$^+$) reactors, reported as a percentage (in the range 0% to 50%) out of the total microbial population, from day 79 until the end of the anaerobic process operation (day 217).

Moerdijk, the Netherlands. In both of these systems, seawater is pumped into the system either for oil recovery or in the cooling towers within the petrochemical plant, likely stimulating salt adaptation of the resident microorganisms. With the taxonomic classification performed using the GTDB (42), *M. harundinacea*, as all the other members of the order *Methanosarcinales*, is classified under the phylum *Halobacterota*. This phylum was proposed in an extensive study of genome phylogenies utilizing a concatenated alignment of 122 marker proteins for archaea (48) and includes the class *Halobacteria*, within which most salt-requiring and salt-resistant archaea are grouped. However, due to the profound changes that the archaeal phylogeny and taxonomy has undergone in recent years (49–52), the adaptation of these *Methanothrix* spp. to salinity could have occurred independently from the origin of their phylogenomic classification. Through the application of FISH on both granule types (28, 32) we found that *Methanothrix* cells were aggregated in two sorts of clusters: fibril-like shaped, with short filaments approaching each other (Fig. 3A); and round shaped, rich in rods (Fig. 3B). We hypothesized that the two shapes of clusters could represent different subspecies of *M. harundinacea*, as was also indicated by the wide range of similarities observed among *Methanothrix*-affiliated 16S rRNA genes detected via clonal analysis on DNA samples extracted from same reactors granules (28). Our findings here suggest that the round and fibril-like clusters could have in fact been comprised of two different *Methanothrix_A* species, MAG_279 and MAG_280. *Methanothrix* spp. indeed tend to aggregate as compact clusters under stress conditions, as previously observed within granules grown at thermophilic temperatures (53, 54). Both aggregation behaviors were particularly visible in HS granules, where the application of FISH combined
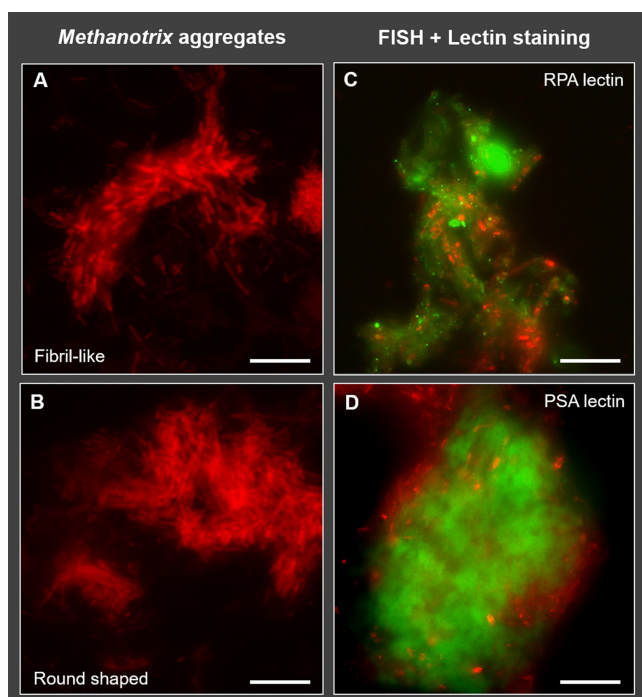
**FIG 3** In (A) and (B), epifluorescence images of the two kind of *Methanotrix* aggregated clusters visualized by FISH with probe ARC915 in HS granules. In (C) and (D), epifluorescence images the same samples stained using RPA (specific for N-Acetyl Galactosamine rich glycosides) and PSA (specific for branched Mannose) lectins combined with probe MSMX860 (*Methanosarcinales*), which proved the association of the two *Methanothrix* shapes with different exopolymeric structures. Size bar is 10 $\mu$m.

with lectin staining highlighted different EPS and glycoconjugate patterns/structures (capsular and cloudy EPS) surrounding these *Methanothrix* clusters (Fig. 3C and D). In addition to this physiological evidence, the genome-resolved metagenomics and functional annotation of the MAGs performed here identified several genetic loci that are potentially involved in the production and modification of surface glycoconjugates, as discussed below.

**Protein glycosylation and outer layer modifications functions.** The production of EPS has been identified in different groups of archaea, predominantly in halophiles and thermophiles (55), but not yet in methanogens. Proteins and carbohydrates in EPS are often combined in the form of glycoconjugates (56). N-glycosylation is one of the most prevalent protein modification processes at the microbial surface layer (S-layer), in which sugars are covalently attached to asparagine residues (57). In O-glycosylation, carbohydrates are instead transferred to serine or threonine residues, but this mechanism has been less studied within the domain *Archaea* (58, 59). Archaea have unique N-glycosylation features in comparison to eukaryotes and bacteria, including the ability to vary the N-glycan composition under different growth conditions (57, 60). For instance, N-glycosylation of S-layer proteins is a salinity response in the archaeon *H. volcanii*, and at least three N-glycan structures have been identified when this microorganism is growing at different salt concentrations (57). While under high salinity stress (3.5 M NaCl), *H. volcanii* was observed to decorate the N-glycans with mannose as the sugar, at lower salinity (1.75 M NaCl) a different pathway was observed, involving rhamnose as the sugar on the glycosylated proteins (39). This ability to change glycosylation patterns likely relies on the presence of multiple copies of the gene for the oligosaccharyl-transferase (OST) AglB, which is responsible for delivering glycan to target protein asparagine residues. The presence of multiple AglB-encoding sequences within a single genome could potentially provide different specificities for different glycan moieties (61). In our previous work, we showed that the EPS surrounding the round-shaped methanogenic cells clusters in the UASB granules
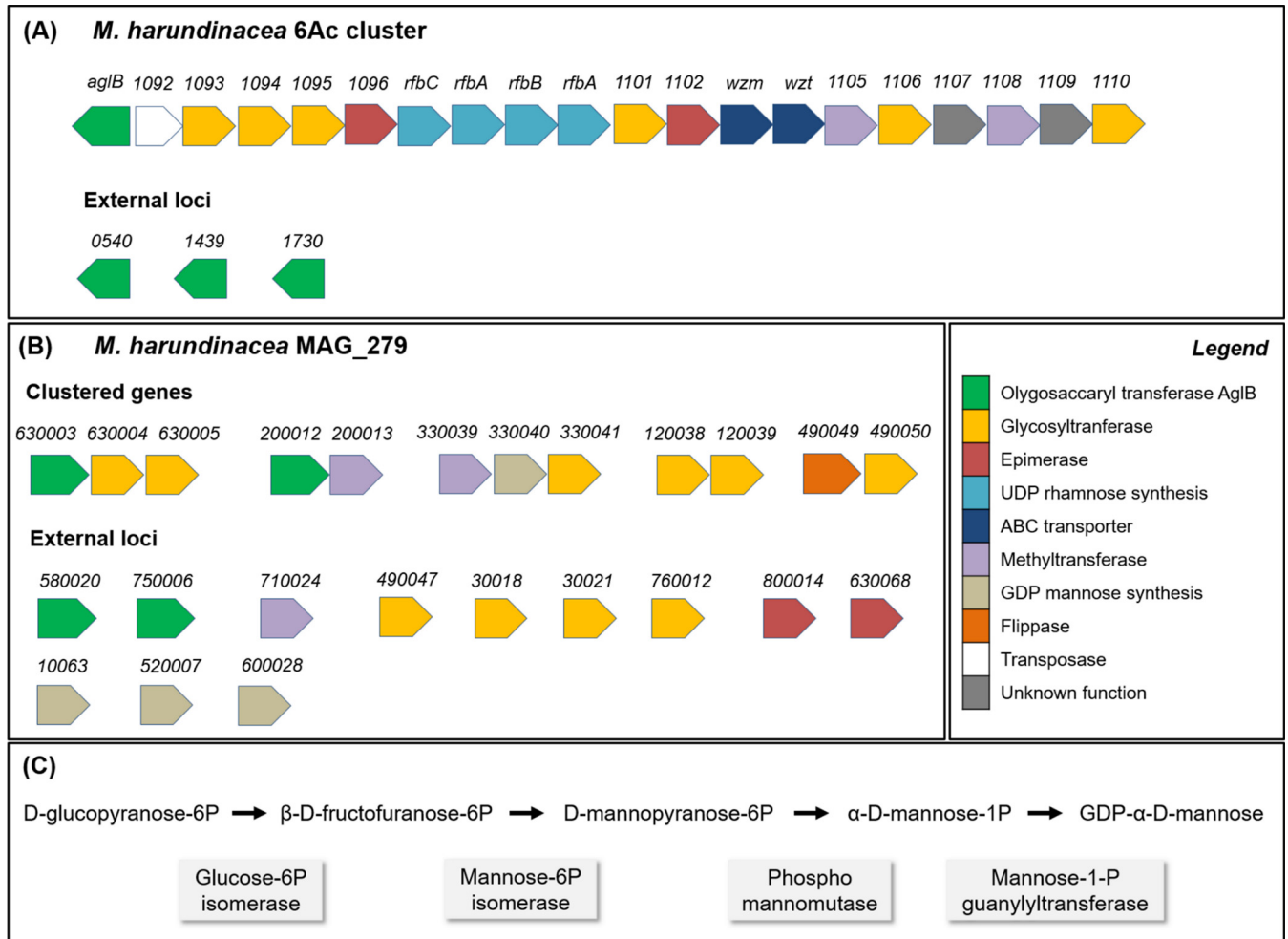
**FIG 4** Schematic representation showing the N-glycosylation genes organization in *M. harundinacea* 6Ac (A) and in *M. harundinacea* MAG_279 (B). In the *M. harundinacea* 6Ac genome, a cluster of 20 genes for the N-glycosylation pathway is present (loci 1091 to 1110), resembling the one of *H. volcanii*. In MAG_279 instead, the putative N-glycosylation genes were grouped into small clusters. In (C), the mannose synthesis pathway as identified in MAG_279.

shifted from a thick outer layer, identified at low salinity, to a cloud-like EPS structure at high salinity (32) (Fig. S4). These EPS shapes were composed of similar sugars, which differed in the arrangement of their glycoconjugate patterns (description in Fig. S4). The cloud-like EPS was found to be rich in branched mannose, of which the likely function was to bind and "inactivate" $Na^+$, shielding the *Methanothrix* cells from excess intake (32, 62). Mannose is the major component of the EPS excreted by many halophilic archaea, such as *Haloferax mediterranei* (63), *H. volcanii* (64), and *Haloarcula japonica* (65), or other extremophiles like *Thermococcus litoralis* (66). In N-glycosilation by *H. volcanii*, several glycosyl transferases (GT) are responsible for sequentially adding carbohydrate residues onto a dolichol phosphate (DolP) carrier on the inside of the cell to build up a glycan. Then, a methyltransferase (MT) catalyzes the addition of a methyl group, a flippase delivers the DolP-bound glycan across the plasma membrane, and finally an OST (AglB) transfers it to the target asparagine residue (67). An *in-silico* analysis of the *M. harundinacea* 6Ac genome revealed the presence of a cluster coding for an N-glycosylation pathway resembling that of *H. volcanii* (32, 39) and multiple copies of the OST AglB, associated with a sub-cluster for rhamnose synthesis (Fig. 4A). In *M. harundinacea* MAG_279, we detected multiple genes encoding for GT, as well as three genes for MT, one gene for a flippase, two genes for epimerase, and four genes encoding for the OST AglB putatively involved in the N-glycosylation (Fig. 4B; Table S2).

By analyzing the four OSTs genes found in *M. harundinacea* MAG_279, we detected

two functional domains able to catalyze either N- or O-glycosilation of membrane proteins. One is an archaeo STT3 (archaeosortase A system-associated) domain, belonging to the subfamily of AglB (Table S3), which catalyzes the transfer of a defined glycan from DolP- to an asparagine residue. The STT3 domain occurs in the phylum *Euryarchaeota*, and particularly within the orders *Archaeoglobales* and *Halobacteriales* (including *H. volcanii*), as well as class II methanogens including the orders *Methanomicrobiales* and *Methanosarcinales* (including *M. harundinacea*) (68). The second domain belongs to the PMT_2 superfamily, which encodes for the dolichyl-phosphate-mannose-protein mannosyltransferase (PMT) that catalyzes the transfer of a mannose from Dol-P-mannose on serine or threonine residues (69) (Table S3). In addition, all of the genes for a putative N-glycosilation pathway were detected in *Methanothrix_A* MAG_280 (Table S2). The three OST genes identified in MAG_280 also encoded for enzymes with STT3 and PMT2 domains, but they were distinct from those of MAG_279 (Table S3), which could also explain the observed differences in their physiologies at high salinity. Moreover, the amino acid identity between the OSTs in *Methanothrix*_A MAGs and other salinity-adapted methanogens (*M. maripaludis* C5 and *M. mazei* Go1) was about 40% (Fig. 1), and such distinct amino acid signature could indicate the potential unique specificity of *Methanothrix*_A OSTs in their salinity-stress related glycosylation pathways (Fig. 1). Methanogens do not assimilate carbohydrates, but rather rely on gluconeogenesis for their formation (e.g., pentoses and hexoses) (70). All three of the *Methanotrix*-related MAGs possessed the genetic repertoire to produce glucose from pyruvate via gluconeogenesis through the same Embden-Meyerhof-Parnas (EMP) pathway as identified in *Methanococcus maripaludis* (Fig. S5) (70, 71). As observed in the archaeon *Pyrococcus horikoshii* OT3 (72), the GDP-mannose required for the synthesis of Dol-P-mannose from glucose-6-phosphate could be produced by MAG_279 and MAG_280 via the sequential action of a glucose-6-phosphate isomerase, a phophomannomutase and a mannose-1-phosphate guanyltransferase (Fig. 4C, Fig. S6).

Overall, the presence of different OST-encoding genes, together with the different types of EPS structures detected via lectin staining surrounding the methanogenic clusters at low and high salinity (Fig. S4), strongly suggests that the genes identified in this study are connected to a molecular mechanism were salinity influences the specificities for a variety of glycan assemblies in *M. harundinacea* MAG_279. The putative N-glycosylation genes identified in *M. harundinacea* MAG_279 are distantly located, clustered as two or three open reading frames, differently than in the genome of *M. harundinacea* 6Ac in which a total of 20 N-glycosylation genes are clustered together (Fig. 4A), as also observed in *H. volcanii* (32, 60). The clustering of N-glycosylation genes together with at least one AglB gene has been shown within the genomes of other archaeal species (60, 61). In *H. volcanii*, it was demonstrated that most of these clustered genes are transcribed coordinately as a single genomic unit (73, 74). However, the lack of clustered genes putatively participating in the N-glycosylation process is also common, such as in thermophilic archaea (57), and this doesn't limit the functionality of the pathway.

**Assessing the potential of the *Methanothrix*-related MAGs for isoprenoids biosynthesis.** The synthesis of isoprenoids, such as carotenoids, and their integration in the cell membranes is a metabolic adaptation in haloarchaea, which similarly to the EPS function outlined in the previous paragraph, can act as a barrier in hypersaline conditions (75). For instance, *H. volcanii* in most cases is pink-red colored, given the production of like bacterioruberin, a $C_{50}$ carotenoid found in the lipid membrane of many other haloarchaea (76). The stress-related accumulation of carotenoids has been well described for microalgae: stress conditions which cause increase in intracellular reactive oxygen species (ROS) can damage to macromolecules such as DNA, lipids, and proteins (77). To scavenge ROS, microalgae synthesize additional carotenoids to form a protective layer that prevents reactive radicals and lipid peroxidation, and thus preserving the cell membrane integrity (78). The precursor of isoprenoids is geranylgeranyl pyrophosphate (GGPP), which in most microorganisms is synthetized via the mevalonate (MVA) pathway through the C-5 intermediate isopentenyl pyrophosphate

(IPP) (40, 76). All isoprenoids like carotenoids are synthesized by the consecutive condensation of the C-5 monomer IPP to its isomer, dimethylallyl diphosphate (DMAPP) (79). IPP and DMAPP isoprenoid precursors are fundamental for the synthesis of isoprene-based alkyl chains constituting the archaeal membrane lipids, characterized by ether linkages to the glycerol moieties (80). These lipids and ether linkages are thought to improve membrane stability, providing an advantage in different extreme environments, such as high salinity (81, 82). The genomes of *M. harundinacea* MAG_279 and *Methanothrix_A* MAG_280 encode all of the genes for the biosynthesis of GGPP via the MVA pathway as found in *H. volcanii*, with the exception of mevalonate-5-phosphate decarboxylase (PMD) (Fig. S7). The gene for PMD is conserved across all haloarchaea, but is not found in most other archaea (83). Indeed, different MVA pathways have been found in other archaea, for instance in *Thermoplasma acidophilus*, *Aeropyrum pernix*, and *Methanocarcina mazei*, where the function of mevalonate-5-phosphate decarboxylase is substituted by other enzymes/reactions (83–85). In the hyperthermophilic archaeon *A. pernix*, two enzymes can replace the PMD function: (i) a putative aconitase catalyzes the dehydration of mevalonate 5-phosphate to trans-anhydromevalonate 5-phosphate; and (ii) an enzyme belonging to the UbiD-decarboxylase family, together with a UbiX-like partner, converts the intermediate into IPP (83). This *Aeropyrum*-type, modified MVA pathway, seems widely distributed among the domain *Archaea*, mostly among anaerobes, probably because it requires less ATP (83, 85). Through a comparison with putative orthologs of the *A. pernix* enzymes contributing to the modified MVA pathway found in *Methanothrix thermophila* (83), we identified four genes that could correspond to the putative aconitase and the UbiD/UbiX functions in *M. harundinacea* 6Ac, *M. harundinacea* MAG_279, and *Methanothrix_A* MAG_280 (Table S4), and reconstructed the pathway (Fig. 5). In MAG_279 and MAG_280, IPP can be converted to DMAPP via an isopentyl diphosphate delta-isomerase, and DMAPP in turn is converted into GGPP via a geranylgeranyl diphosphate synthase (Fig. 5). GGPP is a $C_{20}$ isoprenoid required for carotenoid synthesis: two molecules of GGPP are condensed to phytoene, which is then converted to lycopene, the precursor of many carotenoids such as bacterioruberin (76). The formation of phytoene is catalyzed by phytoene synthase, which is converted to $C_{40}$ lycopene by phytoene desaturase. Finally, lycopene elongase converts lycopene to $C_{50}$ bacterioruberin (46). We did not identify the set of genes encoding for the bacterioruberin synthesis in MAG_279 and MAG_280, and a gene for phytoene synthase was missing. However, in *Methanothrix_A* MAG_279 we identified a homologue of phytoene desaturase (E-value from HMM search: $8^{E-0.5}$), which could potentially lead to lycopene formation. A phytoene desaturase-like function was also found be involved in the synthesis of hydroxyarchaeol, a typical core structure of archaeal membrane lipids uniquely detected in a limited number of lineages of the methanogenic archaea, including *M. soehgenii* (86). Hydroxyarchaeol can further stabilize membranes under extreme conditions, but a direct correlation with salinity levels has not yet been demonstrated (87).

**Pathways encoding for the production of osmolytes.** While the strategies described above prevent the excess accumulation of $Na^+$ intracellularly, the production and accumulation of osmolytes can protect intracellular structures from osmotic stress, maintaining the cell volume. The latter is also hypothesized to be a mechanism employed to abate salinity stress by the microbiome of the HS reactor. Indeed, $N^\varepsilon$-acetyl-$\beta$-lysine was previously identified as the main osmolyte excreted by the granules from HS reactor when subjected to an abrupt decrease in salinity from 20 g $Na^+$/L to 5 g $Na^+$/L (41). Because $N^\varepsilon$-acetyl-$\beta$-lysine is the predominant compatible solute in methanogenic archaea (34), we investigated whether the genomes of *M. harundinacea* MAG_279 and *Methanothrix_A* MAG_280 could encode for its production, together with other osmolytes. Among methanogens, $N^\varepsilon$-acetyl-$\beta$-lysine has been found to be produced by *Methanosarcina* ssp., *Methanogenium cariaci*, and *Methanohalophilus* FDF1 (88–90), but has never been reported for *Methanothrix* spp. In these methanogens, the synthesis of $N^\varepsilon$-acetyl-$\beta$-lysine is initiated by the production of $\alpha$-lysine from aspartate semi-aldehyde and pyruvate via the diaminopimelate (DAP) pathway (88, 89). The diaminopimelate aminotransferases (DAP-related enzyme encoded by the genes *dapA*,
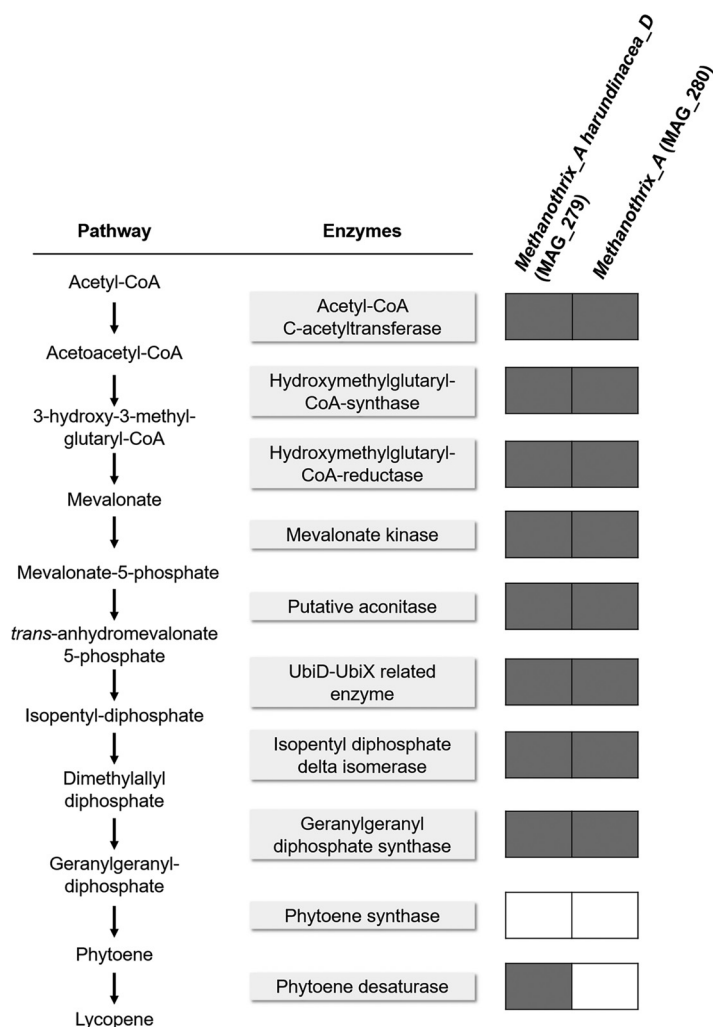
**FIG 5** Putative isoprenoids synthesis via an alternative mevalonate pathway and the annotated enzyme functions found in the two salinity adapted *Methanothrix* MAGs reconstructed in this study (MAG_279 and MAG_280). Gray and white squares indicate the presence or absence of the encoding gene.

*dapB*, *dapL*, and *dapF*), sequentially catalyzes aspartate semi-aldehyde to meso-diamoni-pimelate, that is converted to $\alpha$-lysine by diaminopimelate decarboxylase. Lysine-2, 3-aminomutase (AblA) catalyzes the formation of $\beta$-lysine from $\alpha$-lysine, which is then acetylated to form N$^\varepsilon$-acetyl-$\beta$-lysine by the $\beta$-lysine N(6)-acetyltransferase (AblB) (34, 45, 88). The genomes of *M. harundinacea* MAG_279 and *Methanothrix*_A MAG_280 both encode the DAP pathway (Fig. 6, Fig. S8), as is commonly found in other methanogenic archaea for the synthesis of lysine (91). Homologues of the two genes *ablA* and *ablB*, crucial to synthesize N$^\varepsilon$-acetyl-$\beta$-lysine from $\alpha$-lysine, were detected in both MAG_279 and MAG_280 (Fig. 6, Fig. S8). These genes are typically salt-induced and not active under normal growth conditions (34). Interestingly, the homologue of the *ablB* gene (BLASTP e-value $< 10^{-4}$) observed in *M. harundinacea* MAG_279 and *Methanothrix*_A MAG_280 (Fig. S7) was located in the core gene group of *Methanothrix*_A identified in the pangenome analysis, but it was separately clustered from those found in salinity adapted archaea (Fig. 1; Supplementary Data File 1). Thus, a unique $\beta$-lysine N(6)-acetyltransferase function could be encoded by *Methanothrix*_A MAGs compared with other salinity adapted archaea, as also supported by the low amino acid identity for this *ablB* gene in comparison to the other sequences analyzed (26% for MAG_279 and 45% for MAG_280; Fig. 1).
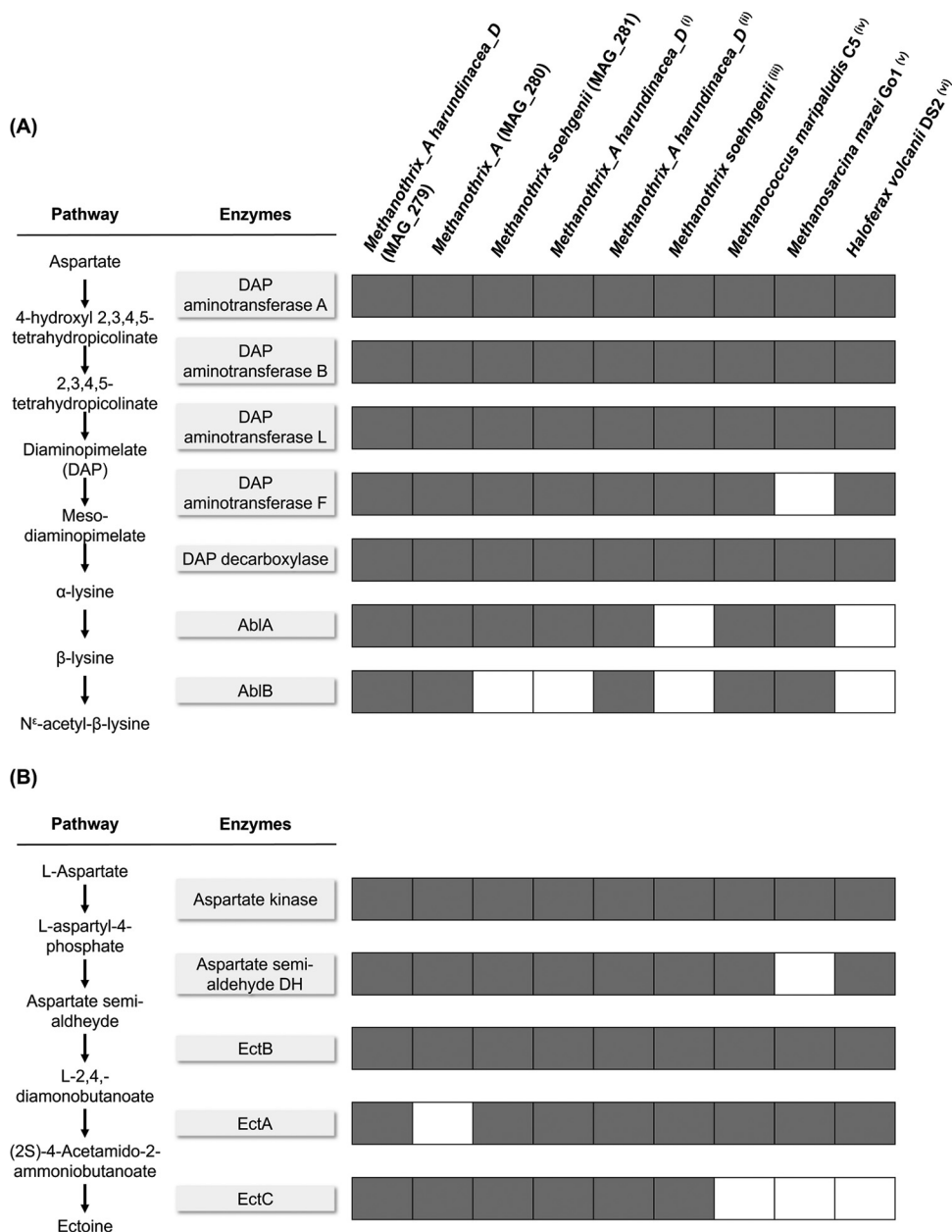
**(A)**



**(B)**



**FIG 6** Putative pathways for the synthesis of the compatible solutes N$^\varepsilon$-acetyl-$\beta$-lysine (in A) and ectoine (in B) and the annotated enzyme functions found in the three *Methanothrix* MAGs reconstructed in this study (MAG_279, MAG_280 and MAG_281) in comparison with other 6 genomes of archaea known to have salinity stress-related functions. The NCBI accession numbers of the genomes used for the comparison are: (i) CP003117; (ii) LGHB00000000.1; (iii) CP002565; (iv) CP000609; (v) AE008384; (vi) CP001956. Per each genome analyzed, gray and white squares indicate the presence or absence of the encoding gene.

Another pathway related to a compatible solute-based osmoprotection mechanism found in the *Methanothrix*-related MAGs was ectoine synthesis. Ectoine is produced from aspartate in a three-step pathway via a L-2,4-diaminobutyrate transaminase (EctB), a 2,4-diaminobutyrate acetyltransferase (EctA), and an ectoine synthase (EctC) (92). As observed for N$^\varepsilon$-acetyl-$\beta$-lysine, where just two of the enzymes of the pathway (AblA and AblB) are crucial for its production, EctC can be regarded as a diagnostic enzyme to identify ectoine producers (93). In a study analyzing 557 genome sequences of archaea to assess the presence of the ectoine biosynthetic genes, *ectC*-type genes, grouped in an *ectABC* gene cluster, have been detected in two species of the *Methanothrix* genus, *M. harundinacea* 6Ac and *M. soehngenii* GP6 (92). In *M. harundinacea* MAG_279 we

identified the three genes *ectA*, *ectB* and *ectC* grouped in a cluster, while in *Methanothrix_A* MAG_280 no homologous of *ectA* were identified (Fig. 6 and Fig. S9). This finding could further explain why *M. harundinacea* MAG_279 had a significant growth advantage over *Methanothrix_A* MAG_280 with the increase of salinity from the LS to the HS reactor.

**Outlooks and conclusive remarks.** The genome-centric metagenomic analysis performed in this study resulted in the identification of a *M. harundinacea* subsp., MAG_279, which appears to be a halotolerant acetoclastic methanogen with the flexibility to function at low (5 g/L Na$^+$) and high salinity conditions (20 g/L Na$^+$). Interestingly, its relative abundance was more than doubled when shifting from 5 to 20 g/L Na$^+$, outcompeting the other dominant methanogen, *Methanothrix_A* MAG_280 for over 200 days of anaerobic reactor operation. Under the conditions described for the UASB reactors in this study, *M. harundinacea* MAG_279 appeared to drive saline granulation by secretion of complex EPS structures, of which the glycoconjugate patterns varied between low and high salinity probably due to the presence of multiple copies of the OST AglB. EPS secretion by *M. harundinacea* has never been reported before, to the best of our knowledge, while the presence of an acyl homoserine lactone (AHL)-based quorum sensing mechanism for this species (14) indicates that it has a role in triggering self-aggregation, and thus granulation. The genomic evidence presented in this study adds support to our previous observations that *Methanothrix* spp. can produce the osmolyte N$^\varepsilon$-acetyl-$\beta$-lysine (41), which is again a novel reported attribute for this genus. N$^\varepsilon$-acetyl-$\beta$-lysine is a rare osmolyte, the synthesis of which can be induced by a salinity increase similar to such conditions that trigger shifts in EPS glycoconjugate composition. Thus, future research is needed on the protein expression patterns activated under the above-mentioned environmental conditions for EPS secretion, glycoconjugate pattern shifts and the production of osmolytes within anaerobic bioreactors. The metabolic potential to synthesize the commercially valuable isoprenoid-precursors as IPP and isoprenoids as GGPP by both MAG_279 and MAG_280 could be seen as an added value when upscaling UASB processes with *Methanothrix*-enriched granules, in addition to their utility in coping with salinity stress. Indeed, in addition to renewable methane generation, anaerobic processes could be potentially utilized as biofactories, where value-added compounds can be produced at significant concentrations and then recovered from excess waste biomass (94). This proposition warrants a more detailed investigation of the economic feasibility and efficiency under different reactor operating configurations. Finally, this improved understanding of the mechanisms in which *Methanothrix* spp. are able to adapt to perform high-efficiency methanisation of organic waste at high salinities opens the door for further valorization of many wastewater streams that have previously been considered problematic.

## MATERIALS AND METHODS

**Bioreactor operation and sample collection.** Two lab-scale UASB reactors were operated for 217 days at a temperature of 35 ± 1°C to treat synthetic wastewater with different salinities: 5 g Na$^+$/L, in the LS reactor, and 20 g Na$^+$/L in the HS reactor. The start-up inoculum was a salinity adapted biomass (~8 g/L Na$^+$) originating from the full-scale UASB reactor of the Shell plant in Moerdijk, the Netherlands, treating acetic/benzoic acid-rich wastewater. Details on the reactor operation are reported in Sudmalis, Gagliano, et al. (29) and in Table S1. Fresh granules were sampled at 10 time points from the sludge bed of the UASBs after granules had formed (Fig. S1), and stored at −20°C until DNA extraction.

**DNA extraction and metagenomic sequencing.** Genomic DNA was extracted from ~500 mg of granules using the FastDNA SPIN kit for soil (MPBio, USA) according to the manufacturer's instructions. Prior the extraction, granules were washed with phosphate-buffered saline (15 min, 37° C), and then pre-treated by sonication (40 kHz, 50 W, 30 s) to disrupt the EPS matrix and facilitate the subsequent cell lysis. After extraction, the DNA concentration and purity were measured with the NanoDrop spectrophotometer (Thermo Fisher Scientific, Germany). Due to residual guanidine salts from the extraction kit, an additional purification procedure was performed. First, one volume of phenol:chloroform:isoamyl alcohol (25:24:1, Sigma-Aldrich, Germany) was mixed with the sample and centrifuged for 15 min at 13,000 × *g*. The recovered supernatant was then treated with RNase (RNase A, Promega, USA) following the manufacturer's protocol. DNA precipitation was performed by adding sodium acetate 3M (1:10 vol/vol) plus two volumes of 70% cold ethanol, and kept at −20° C for 3 h. The recovered DNA pellet was washed with 70% ethanol twice, and then air-dried. Finally, samples where re-suspended in Tris-EDTA (pH 8), and the DNA concentration was measured using the Qubit dsDNA HS assay kit and the Qubit Fluorometer (Thermo

Fisher Scientific, USA) following the manufacturer's instructions. DNA samples were shipped to the DOE Joint Genome Institute (JGI), where they were sequenced using the Illumina NovaSeq platform in paired 151 bp read mode, with an average insert size of 279 bp. The average size of the raw metagenomes was 21 ± 3 Gbp (Table S5).

**Fluorescence microscopy.** Fluorescence *in situ* hybridization (FISH) was carried out on fixed granules collected from both reactors at the end of the reactor operation period by using ARC915 (domain *Archaea*) and MSMX860 (order *Methanosarcinales*) probes to visualize *Methanothrix* cells. Fresh granules collected from LS and HS reactors were used for the lectin staining analysis to visualize EPS structures, by applying 78 different lectins. FISH and lectin staining procedures, and the details of the epifluorescence and confocal laser scanning microscopy (CLSM) analyses are fully described in Gagliano et al. (32).

**Read quality filtering and de novo metagenome assembly.** Raw metagenome reads were quality filtered using BBDuk (v 38.34) (http://bbtools.jgi.doe.gov) to trim adapter sequences and right quality trim reads where the quality score dropped to 0. BBDuk was also used to remove reads that contained four or more "N" bases, had an average quality score less than 3 across the read, or had a minimum length of 33% of the full read length. Reads that mapped with BBMap to masked human, cat, dog, and mouse references at 93% identity, as well as to common microbial contaminants, were removed according to the JGI Metagenome Workflow (95). Quality-filtered interleaved reads for all time point samples from both the HS and LS reactor were separately co-assembled (by reactor) into contigs using MEGAHIT (v1.1.1), with a minimum contig length of 1,000 bp (96). Quality-filtered reads were then mapped to the co-assembled contigs using BowTie2 (v2.3.4.3) (97) with default parameters, and the mapping output was converted to BAM files with SAMtools (v1.9) (98). Anvi'o (v5.5) (99) was used to generate a contigs database, which stores tetranucleotide frequencies for each contig and uses Prodigal (v2.6.3) (100) to identify open reading frames (ORFs). Subsequently, profile databases were created using Anvi'o (v5.5) for each sample, which hosts contig coverage information by parsing the mapping (BAM) files. The sample profile databases were then merged, and the contigs were clustered in Anvi'o using CONCOCT (v1.1.0) (101) to form genome bins. CheckM (v1.0.13) (102) was used to estimate completeness and redundancy of the genomes using single copy genes (SCGs). Bins with more than 75% completion were manually refined using the "*anvi-refine*" option in Anvi'o (v5.5) to generate MAGs. To obtain a dereplicated set of MAGs across both of the reactor sample sets, the refined MAGs from both co-assemblies were dereplicated using dRep (v2.3.2) (103) with a Mash (v2.2) (104) cluster threshold of 0.8 for primary clustering, and the remaining parameters as default values. Taxonomic classification of the dereplicated set of MAGs was performed using the "*classify*" workflow of GTDB-Tk (v0.3.1) (42), using the GTDB taxonomy (release 89) (48). The output results for the archaeal MAGs are reported in Supplementary Data File 2. Quality-filtered reads were then re-mapped to the dereplicated set of MAGs using BowTie2 with default parameters, and the mapping output was converted to BAM files using SAMtools to generate coverage profiles of the dereplicated MAGs across samples for both reactors. The coverage of each dereplicated MAG, and the total cumulative MAG coverage within a metagenome sample, were used to estimate the relative abundance of each MAG as "*coverage of MAG/total cumulative coverage of the sample*." The dereplicated set of MAGs were imported into Anvi'o for phylogenomic tree generation using the programs "*anvi-get-sequences-for-hmm-hits*" and "*anvi-gen-phylogenomic-tree*." These programs used MUSCLE (v3.8.31) (105) to align homologous genes and FastTree (v2.1.10) (106) to generate a Newick tree. Plots for MAG completion and redundancy were generated using the tidyverse package in R (v3.6.3). Phylogenomic trees were plotted using ggtree package (v2.0.2) (107) and heatmaps were in generated using tidyverse package in R.

**Gene annotation and pathway prediction.** The MicroScope annotation platform (108) was used to annotate the ORFs in the recovered MAGs, and the annotations were then used as the input to reconstruct metabolic pathways using Pathway Tools (v23.5) (109) and the BioCyc and MetaCyc databases (110, 111). In Pathway Tools, the pathway prediction algorithm PathoLogic was applied to infer pathways in the annotated MAGs. Further identification of genes that were homologous to closely-related genomes, as well as genomes of organisms with potential salinity tolerance, was performed using BLASTP. Briefly, the amino acid sequences of genes encoding for $N^\varepsilon$-acetyl-$\beta$-lysine and ectoine synthesis and the N-glycosylation pathway were selected from *Methanosarcina mazeii* Go1, *Methanococcus maripaludis* C5, and *Haloferax volcanii* Ds2 representative genomes (from NCBI database). For the mannose synthesis pathway, amino acid sequences for annotated mannose-6P-isomerase genes in *Methanothrix* spp., along with the characterized gene from *Pyrococcus horikoshii* OT3, were obtained from NCBI (Table S6). These reference sequences were used to make individual databases for searches using blast+ (v. 2.11.0), and then the "BLASTP" program was used to query homologous genes within the genomes of interest using an e-value threshold of $10^{-4}$. The positive BLASTP hits for mannose-6P-isomerase genes were further annotated using KofamKOALA (112) and the Pfam protein database (113). Homologous genes encoding for isoprenoid synthesis were searched using hidden Markov models with reference sequences from *Haloferax* members that encoded for phytoene synthase and phytoene desaturase. Briefly, the selected sequences were aligned in a clustal format using MUSCLE (v. 3.8.1551) (105) and a profile HMM was created using HMMER (v3.1b2) (114). The sequences were then searched using this profile HMM via the program "hmmsearch." A threshold value of $10^{-4}$ was used to infer homology to search amino acid-translations of ORFs within the MAGs of interest. Inferred pathways were then manually curated in Pathway Tools to verify predictions from PathoLogic.

**Pangenome analysis.** A pangenome analysis (Fig. 1) was conducted by comparing the three *Methanothrix*-related MAGs with the NCBI genomes of two representative *Methanothrix harundinacea*, *Methanosarcina mazei* Go1, *Methanococcus maripaludis* C5, and *Haloferax volcanii* Ds2, via the anvi'o pangenomics workflow outlined in Delmont and Eren (115). Briefly, an anvi'o genome database ("*anvi-*

*gen-genomes-storage*") was generated to store DNA and amino acid sequences, and functional annotations of the genomes surveyed. The pangenome was computed (using "*anvi-display-pan*") to identify gene clusters based on amino acid sequence similarity, and the anvi'o interface was used to search and highlight genes of interest based on the predicted function obtained via genome annotation.

**Data availability.** NCBI BioProject ID and JGI IMG IDs for the metagenomes are provided in Table S7. All of the archaeal MAGs reported in Fig. 2 were deposited in the NCBI database, and the accession numbers are reported in Table S8.

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.
**SUPPLEMENTAL FILE 1**, XLSX file, 0.02 MB.
**SUPPLEMENTAL FILE 2**, XLSX file, 0.01 MB.
**SUPPLEMENTAL FILE 3**, PDF file, 1.9 MB.

## REFERENCES

1. Smith KS, Ingram-Smith C. 2007. Methanosaeta, the forgotten methanogen? Trends Microbiol 15:150–155. https://doi.org/10.1016/j.tim.2007.02.002.
2. Silva SA, Salvador AF, Cavaleiro AJ, Pereira MA, Stams AJM, Alves MM, Sousa DZ. 2016. Toxicity of long chain fatty acids towards acetate conversion by Methanosaeta concilii and Methanosarcina mazei. Microb Biotechnol 9:514–518. https://doi.org/10.1111/1751-7915.12365.
3. Lee J, Hwang S. 2019. Single and combined inhibition of Methanosaeta concilii by ammonia, sodium ion and hydrogen sulfide. Bioresour Technol 281:401–411. https://doi.org/10.1016/j.biortech.2019.02.106.
4. Magdalena JA, Greses S, González-Fernández C. 2019. Impact of organic loading rate in volatile fatty acids production and population dynamics using microalgae biomass as substrate. Sci Rep 9:18374. https://doi.org/10.1038/s41598-019-54914-4.
5. Chen S, Cheng H, Liu J, Hazen TC, Huang V, He Q. 2017. Unexpected competitiveness of Methanosaeta populations at elevated acetate concentrations in methanogenic treatment of animal wastewater. Appl Microbiol Biotechnol 101:1729–1738. https://doi.org/10.1007/s00253-016-7967-9.
6. Chen S, He Q. 2015. Persistence of Methanosaeta populations in anaerobic digestion during process instability. J Ind Microbiol Biotechnol 42:1129–1137. https://doi.org/10.1007/s10295-015-1632-7.
7. Lü F, Luo C, Shao L, He P. 2016. Biochar alleviates combined stress of ammonium and acids by firstly enriching Methanosaeta and then Methanosarcina. Water Res 90:34–43. https://doi.org/10.1016/j.watres.2015.12.029.
8. Leclerc M, Delgènes JP, Godon JJ. 2004. Diversity of the archaeal community in 44 anaerobic digesters as determined by single strand conformation polymorphism analysis and 16S rDNA sequencing. Environ Microbiol 6:809–819. https://doi.org/10.1111/j.1462-2920.2004.00616.x.
9. van Lier JB, van der Zee FP, Frijters CTMJ, Ersahin ME. 2015. Celebrating 40 years anaerobic sludge bed reactors for industrial wastewater treatment. Rev Environ Sci Biotechnol 14:681–702. https://doi.org/10.1007/s11157-015-9375-5.
10. Wiegant WM. 1988. The spaghetti theory on anaerobic granular sludge formation, or the inevitability of granulation., p 146–152. *In* Lettinga G (ed), Granular anaerobic sludge, Microbiology and technology. Backhuys, the Netherlands.
11. Angenent LT, Sung S, Raskin L. 2004. Formation of granules and Methanosaeta fibres in an anaerobic migrating blanket reactor (AMBR). Environ Microbiol 6:315–322. https://doi.org/10.1111/j.1462-2920.2004.00597.x.
12. Zheng D, Angenent LT, Raskin L. 2006. Monitoring granule formation in anaerobic upflow bioreactors using oligonucleotide hybridization probes. Biotechnol Bioeng 94:458–472. https://doi.org/10.1002/bit.20870.
13. Li L, Zheng M, Ma H, Gong S, Ai G, Liu X, Li J, Wang K, Dong X. 2015. Significant performance enhancement of a UASB reactor by using acyl homoserine lactones to facilitate the long filaments of Methanosaeta harundinacea 6Ac. Appl Microbiol Biotechnol 99:6471–6480. https://doi.org/10.1007/s00253-015-6478-4.
14. Zhang G, Zhang F, Ding G, Li J, Guo X, Zhu J, Zhou L, Cai S, Liu X, Luo Y, Zhang G, Shi W, Dong X. 2012. Acyl homoserine lactone-based quorum sensing in a methanogenic archaeon. ISME J 6:1336–1344. https://doi.org/10.1038/ismej.2011.203.
15. Feng H, Ding Y, Wang M, Zhou G, Zheng X, He H, Zhang X, Shen D, Shentu J. 2014. Where are signal molecules likely to be located in anaerobic granular sludge? Water Res 50:1–9. https://doi.org/10.1016/j.watres.2013.11.021.
16. Beveridge TJ, Patel GB, Harris BJ, Sprott GD. 1986. The ultrastructure of Methanothrix concilii, a mesophilic aceticlastic methanogen. Can J Microbiol 32:703–710. https://doi.org/10.1139/m86-128.
17. Pellerin P, Fournet B, Debeire P. 1990. Evidence for the glycoprotein nature of the cell sheath of Methanosaeta-like cells in the culture of Methanothrix soehngenii strain FE. Can J Microbiol 36:631–636. https://doi.org/10.1139/m90-109.
18. Nomura T. 2012. Control of microbial adhesion using fine particle technology. Advanced Powder Technology 23:532–537. https://doi.org/10.1016/j.apt.2012.05.003.
19. Gagliano MC, Ismail SB, Stams AJM, Plugge CM, Temmink H, Van Lier JB. 2017. Biofilm formation and granule properties in anaerobic digestion at high salinity. Water Res 121:61–71. https://doi.org/10.1016/j.watres.2017.05.016.

20. Muñoz Sierra JD, Oosterkamp MJ, Wang W, Spanjers H, van Lier JB. 2019. Comparative performance of upflow anaerobic sludge blanket reactor and anaerobic membrane bioreactor treating phenolic wastewater: overcoming high salinity. Chemical Engineering J 366:480–490. https://doi.org/10.1016/j.cej.2019.02.097.

21. Foglia A, Akyol Ç, Frison N, Katsou E, Eusebi AL, Fatone F. 2020. Long-term operation of a pilot-scale anaerobic membrane bioreactor (AnMBR) treating high salinity low loaded municipal wastewater in real environment. Separation and Purification Technology 236:116279. https://doi.org/10.1016/j.seppur.2019.116279.

22. Buscio V, López-Grimau V, Álvarez MD, Gutiérrez-Bouzán C. 2019. Reducing the environmental impact of textile industry by reusing residual salts and water: ECUVal system. Chemical Engineering J 373:161–170. https://doi.org/10.1016/j.cej.2019.04.146.

23. Liu X, Dai J, Ng T-L, Chen G. 2019. Evaluation of potential environmental benefits from seawater toilet flushing. Water Res 162:505–515. https://doi.org/10.1016/j.watres.2019.07.016.

24. Hulshoff Pol LW, de Castro Lopes SI, Lettinga G, Lens PNL. 2004. Anaerobic sludge granulation. Water Res 38:1376–1389. https://doi.org/10.1016/j.watres.2003.12.002.

25. Vyrides I. 2015. Anaerobic treatment of organic saline waste/wastewater: overcome salinity inhibition by addition of compatible solutes, p 105–117. *In* Sukla LB, Pradhan N, Panda S, Mishra BK (ed), Environmental microbial biotechnology. Springer International Publishing, Cham, Switzerland.

26. Bruus JH, Nielsen PH, Keiding K. 1992. On the stability of activated sludge flocs with implications to dewatering. Water Res 26:1597–1604. https://doi.org/10.1016/0043-1354(92)90159-2.

27. Vallero MVG, Hulshoff Pol LW, Lettinga G, Lens PNL. 2003. Effect of NaCl on thermophilic (55 degrees C) methanol degradation in sulfate reducing granular sludge reactors. Water Res 37:2269–2280. https://doi.org/10.1016/S0043-1354(03)00024-1.

28. Gagliano MC, Sudmalis D, Pei R, Temmink H, Plugge CM. 2020. Microbial community drivers in anaerobic granulation at high salinity. Front Microbiol 11:235. https://doi.org/10.3389/fmicb.2020.00235.

29. Sudmalis D, Gagliano MC, Pei R, Grolle K, Plugge CM, Rijnaarts HHM, Zeeman G, Temmink H. 2018. Fast anaerobic sludge granulation at elevated salinity. Water Res 128:293–303. https://doi.org/10.1016/j.watres.2017.10.038.

30. Ma K, Liu X, Dong X. 2006. Methanosaeta harundinacea sp. nov., a novel acetate-scavenging methanogen isolated from a UASB reactor. Int J Syst Evol Microbiol 56:127–131. https://doi.org/10.1099/ijs.0.63887-0.

31. Zhu J, Zheng H, Ai G, Zhang G, Liu D, Liu X, Dong X. 2012. The genome characteristics and predicted function of methyl-group oxidation pathway in the obligate aceticlastic methanogens, Methanosaeta spp. PLoS One 7:e36756. https://doi.org/10.1371/journal.pone.0036756.

32. Gagliano MC, Neu TR, Kuhlicke U, Sudmalis D, Temmink H, Plugge CM. 2018. EPS glycoconjugate profiles shift as adaptive response in anaerobic microbial granulation at high salinity. Front Microbiol 9:1423. https://doi.org/10.3389/fmicb.2018.01423.

33. Roesser M, Müller V. 2001. Osmoadaptation in bacteria and archaea: common principles and differences. Environ Microbiol 3:743–754. https://doi.org/10.1046/j.1462-2920.2001.00252.x.

34. Pflüger K, Baumann S, Gottschalk G, Lin W, Santos H, Müller V. 2003. Lysine-2,3-aminomutase and $\beta$-lysine acetyltransferase genes of methanogenic archaea are salt induced and are essential for the biosynthesis of N$\varepsilon$-acetyl-$\beta$-lysine and growth at high salinity. Appl Environ Microbiol 69:6047–6055. https://doi.org/10.1128/AEM.69.10.6047-6055.2003.

35. Decho AW, Gutierrez T. 2017. Microbial extracellular polymeric substances (EPSs) in ocean systems. Front Microbiol 8:922. https://doi.org/10.3389/fmicb.2017.00922.

36. Sandhya V, Ali SZ. 2015. The production of exopolysaccharide by Pseudomonas putida GAP-P45 under various abiotic stress conditions and its role in soil aggregation. Microbiology 84:512–519. https://doi.org/10.1134/S0026261715040153.

37. Philips J, Rabaey K, Lovley DR, Vargas M. 2017. Biofilm formation by Clostridium ljungdahlii is induced by sodium chloride stress: experimental evaluation and transcriptome analysis. PLoS One 12:e0170406. https://doi.org/10.1371/journal.pone.0170406.

38. Nguyen P-T, Nguyen T-T, Vo T-N-T, Nguyen T-T-X, Hoang Q-K, Nguyen H-T. 2021. Response of Lactobacillus plantarum VAL6 to challenges of pH and sodium chloride stresses. Sci Rep 11:1301. https://doi.org/10.1038/s41598-020-80634-1.

39. Guan Z, Naparstek S, Calo D, Eichler J. 2012. Protein glycosylation as an adaptive response in Archaea: growth at different salt concentrations leads to alterations in Haloferax volcanii S-layer glycoprotein N-glycosylation. Environ Microbiol 14:743–753. https://doi.org/10.1111/j.1462-2920.2011.02625.x.

40. Falb M, Müller K, Königsmaier L, Oberwinkler T, Horn P, von Gronau S, Gonzalez O, Pfeiffer F, Bornberg-Bauer E, Oesterhelt D. 2008. Metabolism of halophilic archaea. Extremophiles 12:177–196. https://doi.org/10.1007/s00792-008-0138-x.

41. Sudmalis D, Millah SK, Gagliano MC, Butré CI, Plugge CM, Rijnaarts HHM, Zeeman G, Temmink H. 2018. The potential of osmolytes and their precursors to alleviate osmotic stress of anaerobic granular sludge. Water Res 147:142–151. https://doi.org/10.1016/j.watres.2018.09.059.

42. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2019. GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. Bioinformatics 36:1925–1927. https://doi.org/10.1093/bioinformatics/btz848.

43. Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer K-H, Whitman WB, Euzéby J, Amann R, Rosselló-Móra R. 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. 9. Nat Rev Microbiol 12:635–645. https://doi.org/10.1038/nrmicro3330.

44. Whitman WB, Shieh J, Sohn S, Caras DS, Premachandran U. 1986. Isolation and characterization of 22 mesophilic methanococci. Systematic and Applied Microbiology 7:235–240. https://doi.org/10.1016/S0723-2020(86)80012-1.

45. Pflüger K, Ehrenreich A, Salmon K, Gunsalus RP, Deppenmeier U, Gottschalk G, Müller V. 2007. Identification of genes involved in salt adaptation in the archaeon Methanosarcina mazei Gö1 using genome-wide gene expression profiling. FEMS Microbiol Lett 277:79–89. https://doi.org/10.1111/j.1574-6968.2007.00941.x.

46. Haque RU, Paradisi F, Allers T. 2020. Haloferax volcanii for biotechnology applications: challenges, current state and perspectives. Appl Microbiol Biotechnol 104:1371–1382. https://doi.org/10.1007/s00253-019-10314-2.

47. Hu P, Tom L, Singh A, Thomas BC, Baker BJ, Piceno YM, Andersen GL, Banfield JF. 2016. Genome-resolved metagenomic analysis reveals roles for candidate phyla and other microbial community members in biogeochemical transformations in oil reservoirs. mBio 7:e01669-15–e01615. https://doi.org/10.1128/mBio.01669-15.

48. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, Hugenholtz P. 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol 36:996–1004. https://doi.org/10.1038/nbt.4229.

49. Sorokin DY, Makarova KS, Abbas B, Ferrer M, Golyshin PN, Galinski EA, Ciordia S, Mena MC, Merkel AY, Wolf YI, van Loosdrecht MCM, Koonin EV. 2017. Discovery of extremely halophilic, methyl-reducing euryarchaea provides insights into the evolutionary origin of methanogenesis. 8. Nat Microbiol 2:1–11. https://doi.org/10.1038/nmicrobiol.2017.81.

50. Sorokin DY, Makarova KS, Abbas B, Ferrer M, Golyshin PN, Galinski EA, Ciorda S, Mena MC, Merkel AY, Wolf YI, van Loosdrecht MCM, Koonin EV. 2019. Reply to "Evolutionary placement of Methanonatronarchaeia." Nat Microbiol 4:560–561. https://doi.org/10.1038/s41564-019-0358-0.

51. Aouad M, Borrel G, Brochier-Armanet C, Gribaldo S. 2019. Evolutionary placement of Methanonatronarchaeia. Nat Microbiol 4:558–559. https://doi.org/10.1038/s41564-019-0359-z.

52. Watson T. 2019. The trickster microbes that are shaking up the tree of life. Nature 569:322–324. https://doi.org/10.1038/d41586-019-01496-w.

53. Syutsubo K, Harada H, Ohashi A, Suzuki H. 1997. An effective start-up of thermophilic UASB reactor by seeding mesophilically-grown granular sludge. Water Sci Technol 36:391–398. https://doi.org/10.2166/wst.1997.0615.

54. Sekiguchi Y, Kamagata Y, Nakamura K, Ohashi A, Harada H. 1999. Fluorescence in situ hybridization using 16S rRNA-targeted oligonucleotides reveals localization of methanogens and selected uncultured bacteria in mesophilic and thermophilic sludge granules. Appl Environ Microbiol 65:1280–1288. https://doi.org/10.1128/AEM.65.3.1280-1288.1999.

55. Poli A, Di Donato P, Abbamondi GR, Nicolaus B. 2011. Synthesis, production, and biotechnological applications of exopolysaccharides and polyhydroxyalkanoates by archaea. Archaea 2011:e693253.

56. Görner T, de Donato P, Ameil M-H, Montarges-Pelletier E, Lartiges BS. 2003. Activated sludge exopolymers: separation and identification using size exclusion chromatography and infrared micro-spectroscopy. Water Res 37:2388–2393. https://doi.org/10.1016/S0043-1354(02)00553-5.

57. Jarrell KF, Ding Y, Meyer BH, Albers S-V, Kaminski L, Eichler J. 2014. N-Linked glycosylation in archaea: a structural, functional, and genetic

analysis. Microbiol Mol Biol Rev 78:304–341. https://doi.org/10.1128/MMBR.00052-13.

58. Mescher MF, Strominger JL. 1976. Purification and characterization of a prokaryotic glycoprotein from the cell envelope of Halobacterium salinarium. J Biol Chem 251:2005–2014. https://doi.org/10.1016/S0021-9258(17)33647-5.

59. Sumper M, Berg E, Mengele R, Strobel I. 1990. Primary structure and glycosylation of the S-layer protein of Haloferax volcanii. J Bacteriol 172:7111–7118. https://doi.org/10.1128/jb.172.12.7111-7118.1990.

60. Kaminski L, Eichler J. 2014. Haloferax volcanii N-glycosylation: delineating the pathway of dTDP-rhamnose biosynthesis. PLoS One 9:e97441. https://doi.org/10.1371/journal.pone.0097441.

61. Kaminski L, Lurie-Weinberger MN, Allers T, Gophna U, Eichler J. 2013. Phylogenetic- and genome-derived insight into the evolution of N-glycosylation in archaea. Mol Phylogenet Evol 68:327–339. https://doi.org/10.1016/j.ympev.2013.03.024.

62. Sudmalis D, Mubita TM, Gagliano MC, Dinis E, Zeeman G, Rijnaarts HHM, Temmink H. 2020. Cation exchange membrane behaviour of extracellular polymeric substances (EPS) in salt adapted granular sludge. Water Res 178:115855. https://doi.org/10.1016/j.watres.2020.115855.

63. Antón J, Meseguer I, Rodríguez-Valera F. 1988. Production of an extracellular polysaccharide by Haloferax mediterranei. Appl Environ Microbiol 54:2381–2386. https://doi.org/10.1128/aem.54.10.2381-2386.1988.

64. Chimileski S, Franklin MJ, Papke RT. 2014. Biofilms formed by the archaeon Haloferax volcanii exhibit cellular differentiation and social motility, and facilitate horizontal gene transfer. BMC Biol 12:65. https://doi.org/10.1186/s12915-014-0065-5.

65. Nicolaus B, Lama L, Esposito E, Manca MC, Improta R, Bellitti MR, Duckworth AW, Grant WD, Gambacorta A. 1999. Haloarcula spp. able to biosynthesize exo- and endopolymers. J Industrial Microbiology & Biotechnology 23:489–496. https://doi.org/10.1038/sj.jim.2900738.

66. Rinker KD, Kelly RM. 2000. Effect of carbon and nitrogen sources on growth dynamics and exopolysaccharide production for the hyperthermophilic archaeon Thermococcus litoralis and bacterium Thermotoga maritima. Biotechnol Bioeng 69:537–547. https://doi.org/10.1002/1097-0290(20000905)69:5%3C537::AID-BIT8%3E3.0.CO;2-7.

67. Kaminski L, Naparstek S, Kandiba L, Cohen-Rosenzweig C, Arbiv A, Konrad Z, Eichler J. 2013. Add salt, add sugar: N-glycosylation in Haloferax volcanii. Biochem Soc Trans 41:432–435. https://doi.org/10.1042/BST20120142.

68. Haft DH, Payne SH, Selengut JD. 2012. Archaeosortases and exosortases are widely distributed systems linking membrane transit with posttranslational modification. J Bacteriol 194:36–48. https://doi.org/10.1128/JB.06026-11.

69. Schulze S, Pfeiffer F, Garcia BA, Pohlschroder M. 2021. Comprehensive glycoproteomics shines new light on the complexity and extent of glycosylation in archaea. PLoS Biol 19:e3001277. https://doi.org/10.1371/journal.pbio.3001277.

70. Bräsen C, Esser D, Rauch B, Siebers B. 2014. Carbohydrate metabolism in archaea: current insights into unusual enzymes and pathways and their regulation. Microbiol Mol Biol Rev 78:89–175. https://doi.org/10.1128/MMBR.00041-13.

71. Yu JP, Ladapo J, Whitman WB. 1994. Pathway of glycogen metabolism in Methanococcus maripaludis. J Bacteriol 176:325–332. https://doi.org/10.1128/jb.176.2.325-332.1994.

72. Akutsu J, Zhang Z, Morita R, Kawarabayasi Y. 2015. Identification and characterization of a thermostable bifunctional enzyme with phosphomannose isomerase and sugar-1-phosphate nucleotidylyltransferase activities from a hyperthermophilic archaeon, Pyrococcus horikoshii OT3. Extremophiles 19:1077–1085. https://doi.org/10.1007/s00792-015-0779-5.

73. Yurist-Doutsch S, Abu -Qarn M, Battaglia F, Morris HR, Hitchen PG, Dell A, Eichler J. 2008. aglF, aglG and aglI, novel members of a gene island involved in the N-glycosylation of the Haloferax volcanii S-layer glycoprotein. Mol Microbiol 69:1234–1245. https://doi.org/10.1111/j.1365-2958.2008.06352.x.

74. Yurist-Doutsch S, Magidovich H, Ventura VV, Hitchen PG, Dell A, Eichler J. 2010. N-glycosylation in Archaea: on the coordinated actions of Haloferax volcanii AglF and AglM. Mol Microbiol 75:1047–1058. https://doi.org/10.1111/j.1365-2958.2009.07045.x.

75. Giani M, Garbayo I, Vílchez C, Martínez-Espinosa RM. 2019. Haloarchaeal carotenoids: healthy novel compounds from extreme environments. Mar Drugs 17:524. https://doi.org/10.3390/md17090524.

76. Giani M, Miralles-Robledillo JM, Peiró G, Pire C, Martínez-Espinosa RM. 2020. Deciphering pathways for carotenogenesis in Haloarchaea. Molecules 25:1197. https://doi.org/10.3390/molecules25051197.

77. Ruenwai R, Neiss A, Laoteng K, Vongsangnak W, Dalfard AB, Cheevadhanarak S, Petranovic D, Nielsen J. 2011. Heterologous production of polyunsaturated fatty acids in Saccharomyces cerevisiae causes a global transcriptional response resulting in reduced proteasomal activity and increased oxidative stress. Biotechnol J 6:343–356. https://doi.org/10.1002/biot.201000316.

78. Shi T-Q, Wang L-R, Zhang Z-X, Sun X-M, Huang H. 2020. Stresses as first-line tools for enhancing lipid and carotenoid production in microalgae. Front Bioeng Biotechnol 8:610. https://doi.org/10.3389/fbioe.2020.00610.

79. Kuzuyama T, Seto H. 2012. Two distinct pathways for essential metabolic precursors for isoprenoid biosynthesis. Proc Jpn Acad Ser B Phys Biol Sci 88:41–52. https://doi.org/10.2183/pjab.88.41.

80. Villanueva L, Damsté JSS, Schouten S. 2014. A re-evaluation of the archaeal membrane lipid biosynthetic pathway. Nat Rev Microbiol 12:438–448. https://doi.org/10.1038/nrmicro3260.

81. Kates M. 1993. Biology of halophilic bacteria, Part II. Membrane lipids of extreme halophiles: biosynthesis, function and evolutionary significance. Experientia 49:1027–1036. https://doi.org/10.1007/BF01929909.

82. Vinokur JM, Korman TP, Cao Z, Bowie JU. 2014. Evidence of a novel mevalonate pathway in archaea. Biochemistry 53:4161–4168. https://doi.org/10.1021/bi500566q.

83. Hayakawa H, Motoyama K, Sobue F, Ito T, Kawaide H, Yoshimura T, Hemmi H. 2018. Modified mevalonate pathway of the archaeon Aeropyrum pernix proceeds via trans-anhydromevalonate 5-phosphate. Proc Natl Acad Sci U S A 115:10034–10039. https://doi.org/10.1073/pnas.1809154115.

84. Vinokur JM, Cummins MC, Korman TP, Bowie JU. 2016. An adaptation to life In acid through a novel mevalonate pathway. Sci Rep 6:39737. https://doi.org/10.1038/srep39737.

85. Yoshida R, Yoshimura T, Hemmi H. 2020. Reconstruction of the "archaeal" mevalonate pathway from the methanogenic archaeon Methanosarcina mazei in Escherichia coli cells. Appl Environ Microbiol 86:e02889-19. https://doi.org/10.1128/AEM.02889-19.

86. Mori T, Isobe K, Ogawa T, Yoshimura T, Hemmi H. 2015. A phytoene desaturase homolog gene from the methanogenic archaeon Methanosarcina acetivorans is responsible for hydroxyarchaeol biosynthesis. Biochem Biophys Res Commun 466:186–191. https://doi.org/10.1016/j.bbrc.2015.09.001.

87. Li J, Pancost RD, Naafs BDA, Yang H, Liu D, Gong L, Qiu X, Xie S. 2019. Multiple environmental and ecological controls on archaeal ether lipid distributions in saline ponds. Chem Geol 529:119293. https://doi.org/10.1016/j.chemgeo.2019.119293.

88. Sowers KR, Robertson DE, Noll D, Gunsalus RP, Roberts MF. 1990. N epsilon-acetyl-beta-lysine: an osmolyte synthesized by methanogenic archaebacteria. Proc Natl Acad Sci U S A 87:9083–9087. https://doi.org/10.1073/pnas.87.23.9083.

89. Roberts MF, Lai MC, Gunsalus RP. 1992. Biosynthetic pathways of the osmolytes N epsilon-acetyl-beta-lysine, beta-glutamine, and betaine in Methanohalophilus strain FDF1 suggested by nuclear magnetic resonance analyses. J Bacteriol 174:6688–6693. https://doi.org/10.1128/jb.174.20.6688-6693.1992.

90. Sowers KR, Gunsalus RP. 1995. Halotolerance in Methanosarcina spp.: role of N$\varepsilon$-acetyl-$\beta$-lysine, $\alpha$-glutamate, glycine betaine, and K+ as compatible solutes for osmotic adaptation. Appl Environ Microbiol 61:4382–4388. https://doi.org/10.1128/aem.61.12.4382-4388.1995.

91. Liu Y, White RH, Whitman WB. 2010. Methanococci use the diaminopimelate aminotransferase (DapL) pathway for lysine biosynthesis. J Bacteriol 192:3304–3310. https://doi.org/10.1128/JB.00172-10.

92. Widderich N, Czech L, Elling FJ, Könneke M, Stöveken N, Pittelkow M, Riclea R, Dickschat JS, Heider J, Bremer E. 2016. Strangers in the archaeal world: osmostress-responsive biosynthesis of ectoine and hydroxyectoine by the marine thaumarchaeon Nitrosopumilus maritimus. Environ Microbiol 18:1227–1248. https://doi.org/10.1111/1462-2920.13156.

93. Widderich N, Höppner A, Pittelkow M, Heider J, Smits SHJ, Bremer E. 2014. Biochemical properties of ectoine hydroxylases from extremophiles and their wider taxonomic distribution among microorganisms. PLoS One 9:e93809. https://doi.org/10.1371/journal.pone.0093809.

94. Aldridge J, Carr S, Weber KA, Buan NR. 2021. Anaerobic production of isoprene by engineered Methanosarcina species archaea. Appl Environ Microbiol 87:e02417-20. https://doi.org/10.1128/AEM.02417-20.

95. Clum A, Huntemann M, Bushnell B, Foster B, Foster B, Roux S, Hajek PP, Varghese N, Mukherjee S, Reddy TBK, Daum C, Yoshinaga Y, O'Malley R, Seshadri R, Kyrpides NC, Eloe-Fadrosh EA, Chen I-MA, Copeland A, Ivanova NN. 2021. DOE JGI metagenome workflow. mSystems 6:e00804-20. https://doi.org/10.1128/mSystems.00804-20.

96. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 31:1674–1676. https://doi.org/10.1093/bioinformatics/btv033.

97. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.

98. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

99. Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO. 2015. Anvi'o: an advanced analysis and visualization platform for 'omics data. PeerJ 3:e1319. https://doi.org/10.7717/peerj.1319.

100. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119. https://doi.org/10.1186/1471-2105-11-119.

101. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, Lahti L, Loman NJ, Andersson AF, Quince C. 2014. Binning metagenomic contigs by coverage and composition. 11. Nat Methods 11:1144–1146. https://doi.org/10.1038/nmeth.3103.

102. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055. https://doi.org/10.1101/gr.186072.114.

103. Olm MR, Brown CT, Brooks B, Banfield JF. 2017. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. ISME J 11:2864–2868. https://doi.org/10.1038/ismej.2017.126.

104. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, Phillippy AM. 2016. Mash: fast genome and metagenome distance estimation using MinHash. Genome Biol 17:132. https://doi.org/10.1186/s13059-016-0997-x.

105. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. https://doi.org/10.1093/nar/gkh340.

106. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2 – Approximately maximum-likelihood trees for large alignments. PLoS One 5:e9490. https://doi.org/10.1371/journal.pone.0009490.

107. Yu G, Smith DK, Zhu H, Guan Y, Lam TT-Y. 2017. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol Evol 8:28–36. https://doi.org/10.1111/2041-210X.12628.

108. Vallenet D, Engelen S, Mornico D, Cruveiller S, Fleury L, Lajus A, Rouy Z, Roche D, Salvignol G, Scarpelli C, Médigue C. 2009. MicroScope: a platform for microbial genome annotation and comparative genomics. Database (Oxford) 2009:bap021. https://doi.org/10.1093/database/bap021.

109. Karp PD, Paley S, Romero P. 2002. The Pathway Tools software. Bioinformatics 18:S225–S232. https://doi.org/10.1093/bioinformatics/18.suppl_1.S225.

110. Krieger CJ, Zhang P, Mueller LA, Wang A, Paley S, Arnaud M, Pick J, Rhee SY, Karp PD. 2004. MetaCyc: a multiorganism database of metabolic pathways and enzymes. Nucleic Acids Res 32:D438–D442. https://doi.org/10.1093/nar/gkh100.

111. Karp PD, Billington R, Caspi R, Fulcher CA, Latendresse M, Kothari A, Keseler IM, Krummenacker M, Midford PE, Ong Q, Ong WK, Paley SM, Subhraveti P. 2019. The BioCyc collection of microbial genomes and metabolic pathways. Brief Bioinform 20:1085–1093. https://doi.org/10.1093/bib/bbx085.

112. Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, Ogata H. 2020. KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. Bioinformatics 36:2251–2252. https://doi.org/10.1093/bioinformatics/btz859.

113. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, Tosatto SCE, Paladin L, Raj S, Richardson LJ, Finn RD, Bateman A. 2021. Pfam: The protein families database in 2021. Nucleic Acids Res 49:D412–D419. https://doi.org/10.1093/nar/gkaa913.

114. Wheeler TJ, Eddy SR. 2013. nhmmer: DNA homology search with profile HMMs. Bioinformatics 29:2487–2489. https://doi.org/10.1093/bioinformatics/btt403.

115. Delmont TO, Eren AM. 2018. Linking pangenomes and metagenomes: the Prochlorococcus metapangenome. PeerJ 6:e4320. https://doi.org/10.7717/peerj.4320.