Comparative Transcriptomics Sheds Light on Remodeling of Gene Expression during Diazotrophy in the Thermophilic Methanogen *Methanothermococcus thermolithotrophicus*

Nevena Maslać,[a] Chandni Sidhu,[b] Hanno Teeling,[b] Tristan Wagner[a]

aMicrobial Metabolism Research Group, Max Planck Institute for Marine Microbiology, Bremen, Germany
bDepartment of Molecular Ecology, Max Planck Institute for Marine Microbiology, Bremen, Germany

ABSTRACT Some marine thermophilic methanogens are able to perform energy-consuming nitrogen fixation despite deriving only little energy from hydrogenotrophic methanogenesis. We studied this process in *Methanothermococcus thermolithotrophicus* DSM 2095, a methanogenic archaeon of the order *Methanococcales* that contributes to the nitrogen pool in some marine environments. We successfully grew this archaeon under diazotrophic conditions in both batch and fermenter cultures, reaching the highest cell density reported so far. Diazotrophic growth depended strictly on molybdenum and, in contrast to other diazotrophs, was not inhibited by tungstate or vanadium. This suggests an elaborate control of metal uptake and a specific metal recognition system for the insertion into the nitrogenase cofactor. Differential transcriptomics of *M. thermolithotrophicus* grown under diazotrophic conditions with ammonium-fed cultures as controls revealed upregulation of the nitrogenase machinery, including chaperones, regulators, and molybdate importers, as well as simultaneous upregulation of an ammonium transporter and a putative pathway for nitrate and nitrite utilization. The organism thus employs multiple synergistic strategies for uptake of nitrogen nutrients during the early exponential growth phase without altering transcription levels for genes involved in methanogenesis. As a counterpart, genes coding for transcription and translation processes were downregulated, highlighting the maintenance of an intricate metabolic balance to deal with energy constraints and nutrient limitations imposed by diazotrophy. This switch in the metabolic balance included unexpected processes, such as upregulation of the CRISPR-Cas system, probably caused by drastic changes in transcription levels of putative mobile and virus-like elements.

IMPORTANCE The thermophilic anaerobic archaeon *M. thermolithotrophicus* is a particularly suitable model organism to study the coupling of methanogenesis to diazotrophy. Likewise, its capability of simultaneously reducing N_2 and CO_2 into NH_3 and CH_4 with H_2 makes it a viable target for biofuel production. We optimized *M. thermolithotrophicus* cultivation, resulting in considerably higher cell yields and enabling the successful establishment of N_2-fixing bioreactors. Improved understanding of the N_2 fixation process would provide novel insights into metabolic adaptations that allow this energy-limited extremophile to thrive under diazotrophy, for instance, by investigating its physiology and uncharacterized nitrogenase. We demonstrated that diazotrophic growth of *M. thermolithotrophicus* is exclusively dependent on molybdenum, and complementary transcriptomics corroborated the expression of the molybdenum nitrogenase system. Further analyses of differentially expressed genes during diazotrophy across three cultivation time points revealed insights into the response to nitrogen limitation and the coordination of core metabolic processes.

Methanogenic archaea generate about 1 gigaton of methane per year (1). This amounts to about half of the greenhouse gas methane in our atmosphere (2, 3). Methanogenesis is a strictly anaerobic process occurring in habitats in which electron acceptors other than CO_2 are depleted (4). It is accepted that under natural conditions methanogenesis provides an extremely low energy yield, causing methanogens to thrive close to the thermodynamic limits of life (1, 5, 6). Therefore, it came as a surprise when in 1984 two studies proved that *Methanothermococcus thermolithotrophicus* (7) and *Methanosarcina barkeri* (8) can perform nitrogen fixation, a very energy-demanding metabolic process that requires the hydrolysis of at least 16 ATP per molecule of fixed N$_2$ (9).

The *nifH* gene is widely used as a marker to identify nitrogen-fixing organisms. Multiple environmental studies have detected *nifH* genes and transcripts from methanogens in diverse anoxic habitats, such as deep seawater and hydrothermal vent fluids (10), oligotrophic open seas (11), deep-sea methane seep sediments (12), and N$_2$-limited soils of salt marshes (13). Collectively these findings show that methanogens are indeed actively fixing N$_2$ in nature and thereby contributing considerably to global nitrogen cycling. They also imply that methanogens and anaerobic methanotrophs (14) can overcome the largest activation barrier in biology, $+251$ kJ mol^{-1} (15), to break the N$_2$ triple bond.

The reduction of N$_2$ to NH$_3$ is catalyzed by the nitrogenase enzyme complex. The overall organization of this complex is highly conserved among *Bacteria* and *Archaea*. It is composed of a dinitrogenase reductase (iron protein, NifH) and dinitrogenase (iron-molybdenum protein, NifDK) containing one [MoFe$_7$S$_9$C-(*R*)-homocitrate] iron-molybdenum cofactor (FeMo-co) and an [8Fe-7S] P-cluster (15). The nitrogenase-encoding *nif* genes and a plethora of accessory proteins with roles in regulation, biosynthesis, and maturation of metal cofactors can be clustered within one or in several operons or regulons, depending on the organism (15–18). In addition to Nif, which represents the most widespread and well-studied system, at least two alternative nitrogenases are known: vanadium nitrogenase (VFe protein; Vnf) and iron-only nitrogenase (FeFe protein, Anf) (15). Both are considered to be evolutionarily related to the Nif system (19).

Phylogenetic analyses suggest that the ancestral nitrogenase originated in anaerobic methanogenic archaea, from where it was subsequently transferred into the bacterial domain, most probably initially to the *Firmicutes* (15, 19–22). Despite the shared origin and a minimal conserved set of genes required for functionality, archaeal nitrogenases are very distinct from their bacterial homologs. This is reflected by several features: (i) the organization of nitrogenase operon(s) (see Fig. S1 in the supplemental material), (ii) transcriptional regulation by the repressor NrpR and the activator NrpA, and (iii) the unique mode of posttranslational regulation through direct protein-protein interaction (23–25). In the latter case, the inhibition of nitrogenase activity is achieved by NifI$_{1,2}$, a P$_{II}$ family regulator which blocks the NifH binding site on NifDK. Both NifI$_1$ and NifI$_2$ are required for the switch-off (26), which is reversed by 2-oxoglutarate magnesium, and ATP addition. These differences have been extensively studied in the mesophilic methanogen *Methanococcus maripaludis* (16, 17, 26–30) and different *Methanosarcina* strains (31–36) using well-established genetic systems for both.

In this context, it is noteworthy that *M. maripaludis* features a single operon consisting exclusively of *nif* genes (16, 28) (Fig. S1), whereas some *Methanosarcinales* harbor additional *vnf* and *anf* operons coding for V-only and Fe-only nitrogenases, respectively. However, these operons all share the core organization of the *nif* operon, in which *nifH*, *nifD*, and *nifK* genes code for the two structural subunits of the nitrogenase system while *nifE* and *nifN* play a putative role in the synthesis of FeMo-co. Genes *nifI$_1$* and *nifI$_2$* encode P$_{II}$ family regulators that are conserved in archaea, while *nifX* has no homology with the homonymous bacterial gene (Fig. S1). NifX, absent in *Methanosarcinales*, has an unknown function. It is not required for nitrogen fixation in *M. maripaludis*, since *nifX* in-frame deletion mutants are still capable of diazotrophic growth (17).

Only a few studies exist on the diazotrophic physiology of (hyper)thermophilic *Methanococcales*. The N$_2$ fixation capabilities of these hydrogenotrophic methanogens are particularly interesting for several reasons: (i) their nitrogenases form a separate evolutionary branch (19) (Fig. S2 and Fig. S3), (ii) they can perform N$_2$ fixation under very high H$_2$ partial
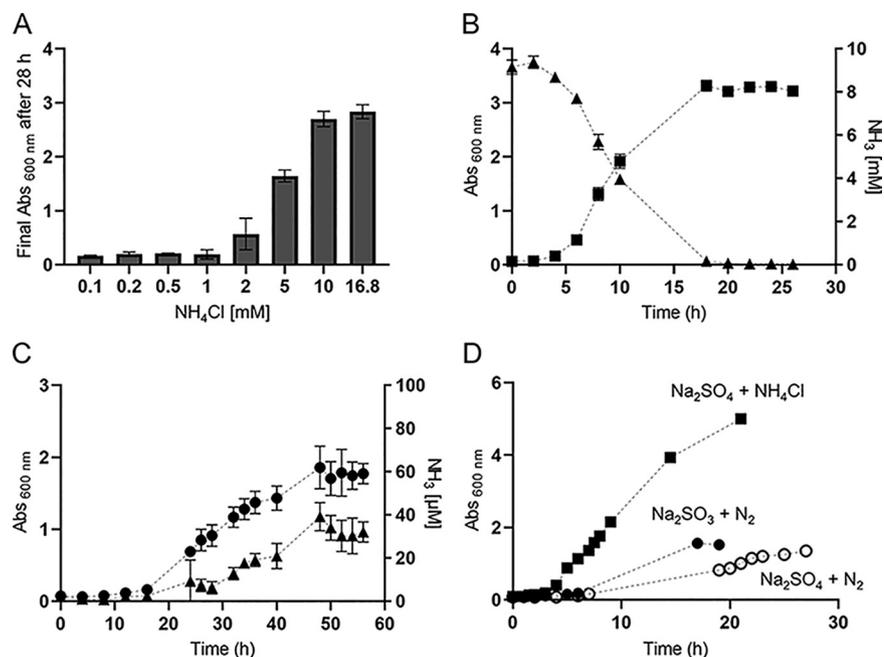
FIG 1 Growth of *Methanothermococcus thermolithotrophicus* on different nitrogen sources. (A) Final A_{600} of *M. thermolithotrophicus* cultures after 28 h of incubation with different NH_4Cl concentrations. (B) Growth curve of *M. thermolithotrophicus* cultures grown on 10 mM NH_4Cl (squares) and NH_4Cl consumption during growth (triangles). (C) Growth curve of diazotrophic *M. thermolithotrophicus* cultures grown on N_2 as the sole nitrogen source (circles) and NH_3 release during the growth (triangles). (D) Growth curves of nondiazotrophic *M. thermolithotrophicus* on 16.8 mM NH_4Cl (squares), a diazotrophic culture grown on Na_2SO_3 (full circles), and a diazotrophic culture grown on Na_2SO_4 (empty circles) in a fermenter. Measurements for panel A were performed in duplicates; measurements for panels B and C were performed in triplicates.

pressure that usually inhibits nitrogenases (7, 8, 37), and (iii) they can fix nitrogen at the highest temperature ever described so far (i.e., up to 92°C) (37). Together, findings suggest a unique adaptation that allows N_2 fixation to operate at both high temperatures (7) and extreme energy limitations.

We investigated the integration of N_2 fixation with other metabolic processes in thermophilic *M. thermolithotrophicus* by using a combined approach of physiological tests and differential transcriptomics. Our results provide new insights into the adaptive strategies of *M. thermolithotrophicus* to energy and nutrient limitation stresses inflicted by Mo-dependent diazotrophy.

RESULTS

Nitrogen acquisition by *M. thermolithotrophicus*. We selected *M. thermolithotrophicus* DSM 2095 due to its remarkable chemolithoautotrophic capabilities and fast growth at 65°C in mineral medium (38, 39). Series of incubations with different NH_4Cl concentrations, ranging from 0.1 to 16.8 mM, showed that *M. thermolithotrophicus* required a minimum of 10 mM NH_4Cl for best growth (Fig. 1A) in our optimized medium (see Materials and Methods for composition) (39, 40). Higher NH_4Cl concentrations of 25 to 200 mM NH_4Cl did not result in higher cell yields, and a 500 mM excess of NH_4Cl led to a decrease (see Fig. S4 in the supplemental material). Continuous monitoring of ammonia consumption during growth on 10 mM NH_4Cl over a span of 26 h confirmed depletion proportional to the observed increase in biomass (Fig. 1B). The culture reached stationary phase after 18 h, during which most of the ammonia was consumed beyond the reliable detection limit.

M. thermolithotrophicus was adapted to diazotrophic conditions after three successive transfers to NH_4Cl-free medium. Under this condition, dissolved N_2 was the sole available nitrogen source. The cell density of the diazotrophic culture was 6 times higher than that reported by Belay and colleagues in their original study on the discovery of diazotrophy in methanogens in 1984 (e.g., Absorbance 600nm 1.85 achieved in 48 h, compared to an
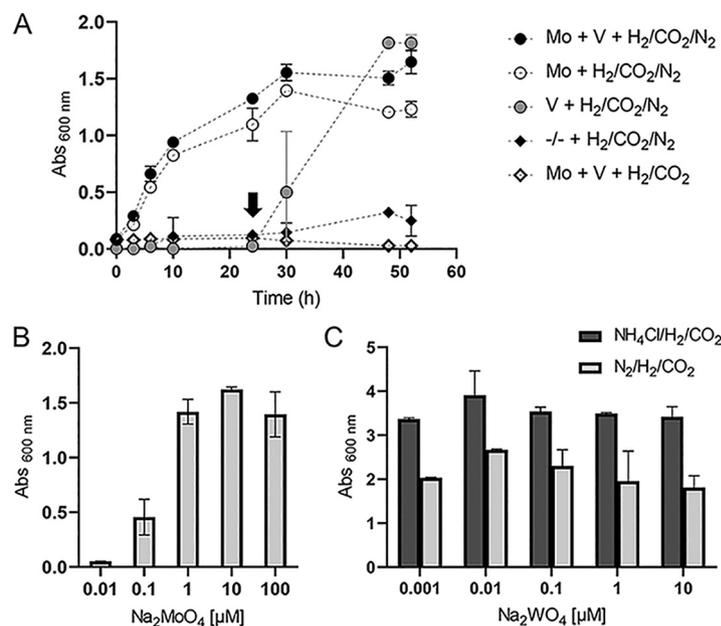
FIG 2 Influence of trace metal availability on diazotrophic growth of *M. thermolithotrophicus*. (A) Growth curves of diazotrophic *M. thermolithotrophicus* grown in the medium with both Mo and V (full circles), without V (empty circles), without Mo (gray circles), and without both Mo and V (full rhomboids) and of the negative control (empty rhomboids). The arrow indicates the supplementation of the culture containing V without Mo with 16.8 mM NH_4Cl. (B) Final A_{600} of *M. thermolithotrophicus* as a function of MoO_4^{2+} concentration. (C) Final A_{600} of nondiazotrophic (dark gray bars) and diazotrophic (light gray bars) *M. thermolithotrophicus* as a function of WO_4^{2+} concentration. A concentration of 10 μM Na_2MoO_4 was used for this experiment. Measurements for panel A were performed in triplicates; measurements for panels B and C were performed in duplicates.

Absorbance 600nm of 0.4 achieved in ∼20 h [7]). Ammonia possibly released in the medium was measured during diazotrophic growth (Fig. 1C), and traces were detected during the mid-exponential phase, reaching a maximum of 40 μM ammonia in the stationary phase (Fig. 1C). However, we cannot exclude that detected ammonia also originated from cell lysis rather than from active excretion.

In addition to batch diazotrophic cultures, *M. thermolithotrophicus* was successfully grown in a 10-liter fermenter continuously supplied with H_2/CO_2 and N_2. To maintain the sulfur source in the medium, we replaced Na_2S used in batch cultures (which would be flushed out as H_2S) with Na_2SO_3 or Na_2SO_4 (see Materials and Methods) (39). Diazotrophic growth was not affected by HSO_3^-, a known inhibitor of methanogenesis (Fig. 1D) (41). While the final cell yield was similar to the one observed in batch cultures, division times were shorter (Fig. 1D). Fermenter-grown cultures supplemented with NH_4Cl had higher final yields than diazotrophic fermenter-grown cultures. However, in comparison to NH_4^+-grown batch cultures, the difference was not very pronounced, as final yields were only slightly higher and the division times were similar.

N_2 fixation was molybdenum dependent. We then investigated the nitrogenase type used for diazotrophy, taking into account that the *M. thermolithotrophicus* genome features only one *nifDK*. It must be noted that *M. thermolithotrophicus* resembles *M. maripaludis* in terms of physiology, and the latter has been shown to require molybdenum (Mo) for its diazotrophic growth (28). Therefore, we first tested for Mo dependency. Growth of cultures lacking either Mo, vanadium (V), or both metals in the medium were monitored simultaneously. Depletion of Mo and V from the medium was achieved by three successive culture transfers to the same medium without Mo and V. The highest optical density at 600 nm (OD_{600}) was reached when *M. thermolithotrophicus* was grown in the control medium containing both Mo and V. Growth was similar in the diazotrophic culture incubated without V, but with a lower final OD_{600} (Fig. 2A). This trend was reproducible and suggested that V acts as a potential growth stimulator under such conditions; however, the potential mechanism remains unclear. No growth was observed in the absence of Mo, with or without V (Fig. 2A).

The minimal required Mo concentration of 0.1 μM was determined in a separate series of incubations with Mo concentrations ranging from 0.01 to 100 μM (Fig. 2B). To check if Mo was essential for methanogenesis (42) or metabolic processes other than nitrogen fixation, we supplemented the culture grown with V in the absence of Mo with NH$_4$Cl, which restored growth after an overnight incubation (Fig. 2A).

An inhibitory effect of tungstate (WO$_4^{2-}$) on diazotrophic growth has already been observed in *Azotobacter vinelandii* OP (43), *Methanosarcina barkeri* 227 (32), and *M. maripaludis* strain S2 (28). In this case, WO$_4^{2-}$ can inhibit MoO$_4^{2-}$ uptake (44), and it can be incorporated into the nitrogenase, thereby rendering the protein inactive (43, 45, 46). As shown by Siemann and co-workers in their work on *Rhodobacter capsulatus*, a WFe nitrogenase does not exhibit N$_2$ or C$_2$H$_2$ reducing activity but can reduce protons (46). These inhibition mechanisms are competitive and thus dependent on the MoO$_4^{2-}$/WO$_4^{2-}$ ratio. Surprisingly, there was no observable inhibitory effect in *M. thermolithotrophicus* (Fig. 2C), even at a 1:1 MoO$_4^{2-}$/WO$_4^{2-}$ ratio. In contrast, diazotrophic growth of *M. maripaludis* was inhibited at a 10:1 MoO$_4^{2-}$/WO$_4^{2-}$ ratio (28). This observation might be explained by different affinities of the WO$_4^{2-}$ and MoO$_4^{2-}$ transport systems in both species. The ModABC transporter (47, 48) is present in both organisms and should transport MoO$_4^{2-}$ and WO$_4^{2-}$ (49). In addition, *M. thermolithotrophicus* has a highly specific WO$_4^{2-}$ transporter, TupABC (50), that is lacking in *M. maripaludis*. Instead, *M. maripaludis* has a third type of tungstate transporter, WtpABC (49, 51), which can transport WO$_4^{2-}$ and MoO$_4^{2-}$ but has a higher affinity for WO$_4^{2-}$ than has been shown for Mod and Tup in *Pyrococcus furiosus* (51). Furthermore, it is also known that Mod transporters can have different affinities for both oxyanions in different organisms (49, 52, 53), suggesting that both transporter specificities and the specificity of the FeMo-co insertion machinery contribute to the W tolerance during diazotrophy.

Diazotrophy shifted expression of a large number of genes. Transcriptomics and DNA microarrays have been used to investigate the complex metabolic and regulatory networks that control N$_2$ fixation in two model organisms: *Azotobacter vinelandii* (54) and *Methanosarcina mazei* Gö1 (36, 55, 56). These studies led to the discovery of three gene clusters (*rnf1*, *rnf2*, and *fix*) coding for electron transfer systems that provide reducing equivalents to the nitrogenase in *A. vinelandii*, as well as multiple potential transcriptional and small RNA (sRNA) regulators in *Methanosarcina mazei* Gö1. Here, we used comparative transcriptome profiling of N$_2$-fixing *M. thermolithotrophicus* cultures versus NH$_4^+$-grown cultures to investigate the metabolic adaptations induced by nitrogen fixation. The experiment was conducted in biological triplicates at three different time points: 3 h (early exponential phase), 21 h (stationary-phase starvation due to the exhaustion of the gas phase), and 25 h postinoculation (see Materials and Methods) (Fig. 3A and Table S1). The gas phase was exchanged after sampling at the 21-h time point, since the cultures consumed H$_2$/CO$_2$. The robustness of our sample data was reflected in a corresponding principal-component analysis (PCA), in which triplicates from different conditions and time points were clustered (Fig. 3B). More importantly, it revealed that 92% of the variation could be explained by the different sample treatments.

Transcriptome profiling across the three time points revealed prominent changes (Fig. 3C). Of the total 1,751 predicted genes (open reading frames [ORFs]), 1,737 genes were found to have nonzero read counts. At 3 h, 12.3% (214) were differentially expressed, with most (122/214) being upregulated under diazotrophic conditions. More than half of these genes (70/122) were of an unknown function without any homologs in the databases (their probable role is discussed further below). The difference was smaller at 21 h, with only 3.3% (57) of the total expressed genes having differential expression; however, at 25 h 17.4% (302) were differentially expressed, with most of them downregulated (257/302) under diazotrophic conditions (Table S2).

To check the overall expression of genes irrespective of differential expression, we mapped reads to predicted genes in terms of transcripts per million (TPM). Details of this analysis are summarized in the sections below.

Nitrogen and molybdate acquisition genes were mostly affected during the early exponential phase. As anticipated, upregulated genes at 3 h included the *nif* operon (*nifHI$_1$I$_2$DKXEN*) and molybdate-acquisition genes (*modABC*). This illustrated a swift response of *M. thermolithotrophicus* to diazotrophic conditions. The highest log$_2$ fold change (FC) was
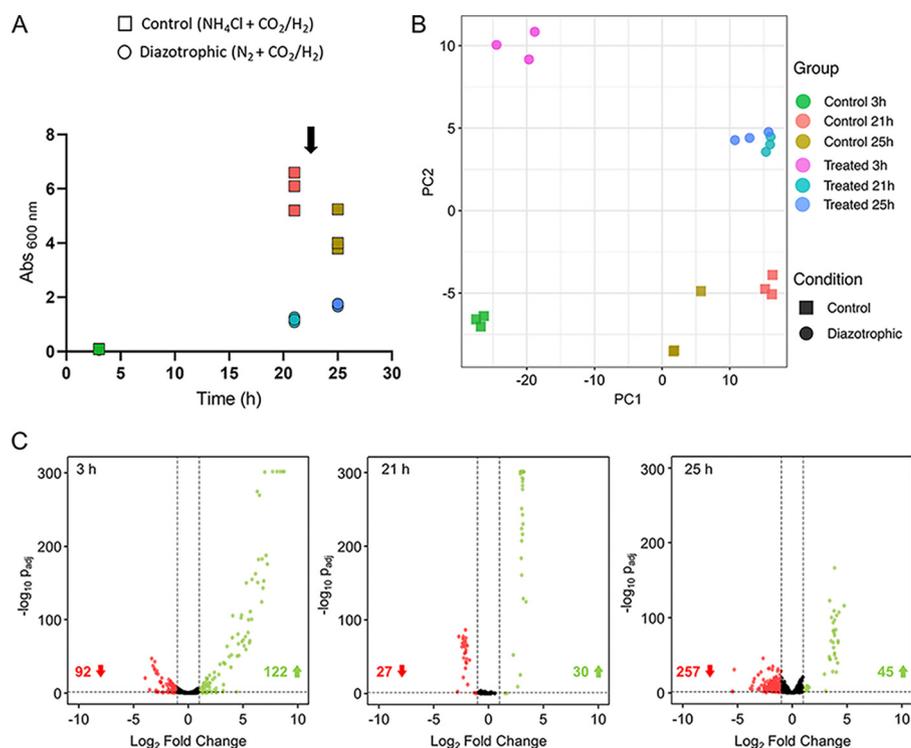
FIG 3 Differential gene expression between NH$_4$Cl-grown and diazotrophic cultures. (A) Points during growth of the control (squares) and diazotrophic (circles) cultures at which samples for transcriptomic profiling were taken (each of the triplicates is shown separately). The arrow indicates the complete gas-phase exchange of the cultures after the sample at 21 h was taken. (B) PCA plot showing the first two principal components that explain the variability in the data using the regularized log count data. Control corresponds to the NH$_4$Cl-grown cultures. (C) Volcano plots of differentially expressed genes in the diazotrophic culture versus control at different time points. Negative and positive log$_2$ fold changes of >1 with adjusted P of <0.05 are shown in red and green, respectively. Values in red and green indicate the number of down- and upregulated genes, respectively.

observed for *nifI$_1$I$_2$HDKX* genes, with values ranging from 8.8 to 7.0 (Table S2). The proteins of the FeMo-co biosynthetic machinery, *nifE* and *nifN*, were also among the upregulated genes, with log$_2$FC of 5.3 and 4.6, respectively (Fig. 4A and B). An essential gene for FeMo-co biosynthesis, *nifB* (57, 58), was transcribed at a lower level, with a log$_2$FC of 1.8.

The *mod* operon, coding for the three subunits of a molybdate ABC transporter (*modABC*), was also upregulated (Fig. 4A), with *modA* being strongly expressed with a log$_2$FC of 5.5 (Table S2). This agreed with our physiological data, that showed a clear Mo dependency under diazotrophic conditions. Although three *modA* instances are present in the genome, only the *modA* gene in the vicinity of the *nif* operon, which is part of the complete *mod* operon, was highly expressed. The *tupA* gene coding for the tungstate transporter was also upregulated (log$_2$FC, 1.7), strengthening the hypothesis that both transporter specificities and adaptation of the cofactor biosynthetic machinery contribute to the W tolerance during diazotrophy.

The other genes upregulated at 3 h included the ammonium transporter *amtB$_2$* and its P$_{II}$ family regulatory protein *glnK$_2$*, with the log$_2$FCs of 6.2 and 5.8, respectively (Table S2). The genome of *M. thermolithotrophicus* harbors two different *amtB* genes with their associated *glnK* regulators (40). Interestingly, *amtB$_1$* and *glnK$_1$* were not differentially expressed (Table S2; Fig. 4A and B), which might point to the existence of an internal regulator or an additional promoter. In addition to the *nif* and *amtB/glnK* operons, which are known to be upregulated during diazotrophy from previous studies (54, 55, 59), the ammonium-assimilating glutamine synthetase (*glnA*) gene was also upregulated (Table S2). However, the transcript level of the second enzyme involved in N-assimilation, glutamate synthase, remained unchanged, suggesting that GlnA is the rate-limiting step in ammonium assimilation.

Unexpectedly, *narK* coding for a putative nitrate transporter and *narB* coding for a molybdopterin-dependent nitrate reductase (60) were also highly expressed during
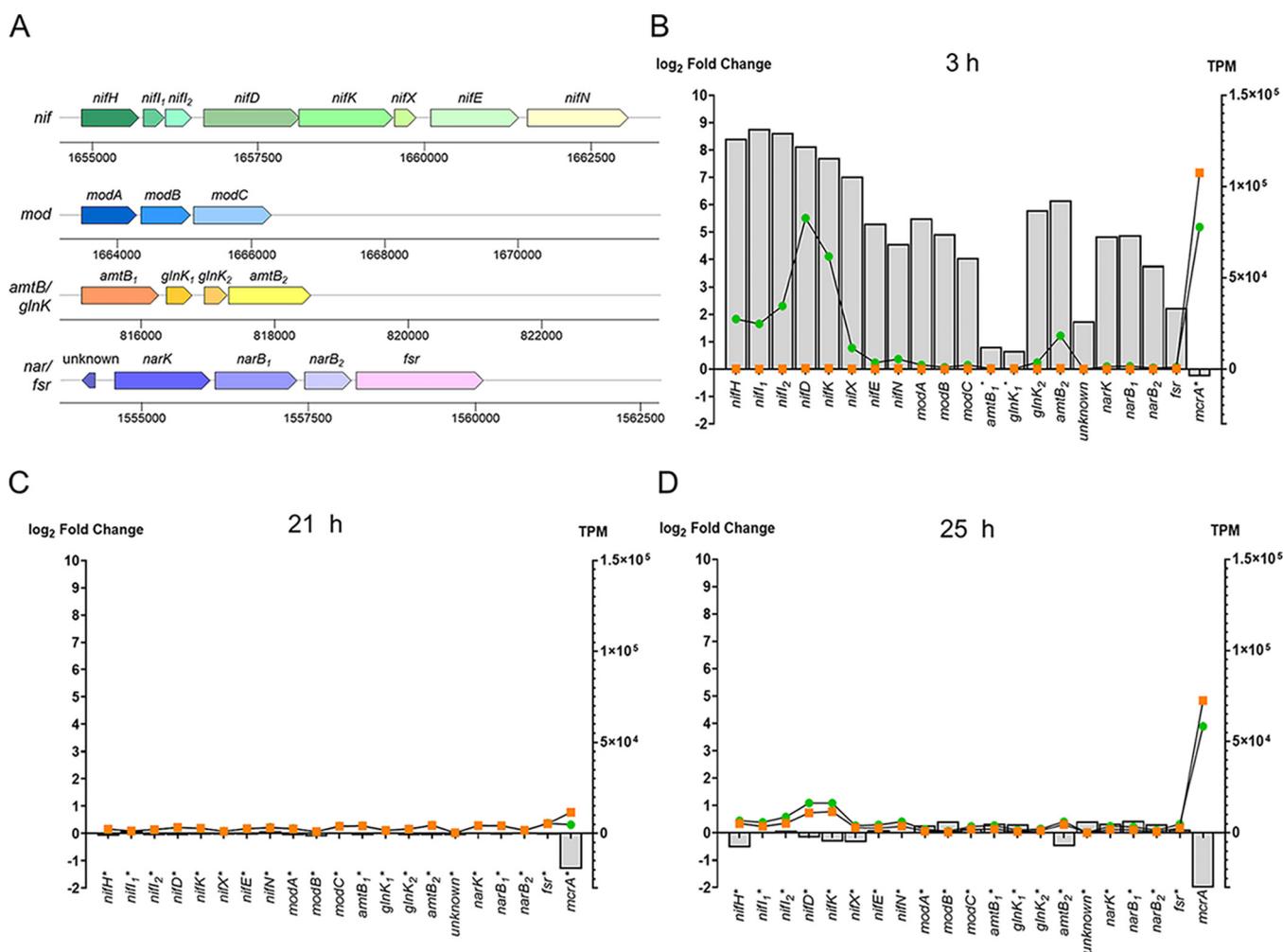
FIG 4 Changes in the expression over time of selected genes involved in nitrogen acquisition. (A) Arrangement of the *nif*, *mod*, *amtB/glnK*, and *nar/fsr* operons in *M. thermolithotrophicus*. (B to D) Log$_2$ fold changes (light gray bars) and TPM of the controls (orange squares) and diazotrophic cultures (green circles) of nitrogen acquisition genes, compared to the *mcrA* gene as reference, after 3 h (B), 21 h (C), and 25 h (D). Genes that were not differentially expressed at a given time point are marked by an asterisk.

diazotrophy in the early exponential stage (Fig. 4A and B; Table S2). If nitrate were imported and reduced in the cell, oxidant nitrite would be generated and could damage the highly oxidation-sensitive methanogenic machinery (61). A gene coding for a F$_{420}$-sulfite reductase isoform (also belonging to Fsr group 1 but different from the one naturally expressed under sulfite conditions [39]), which cooccurs with *narK* and *narB*, would be a plausible candidate for nitrite detoxification, since Fsr has been recognized to catalyze nitrite reduction (39). *M. thermolithotrophicus* has been reported to be able to grow on nitrate as the only source of nitrogen (62), which might involve these genes.

In addition, three genes coding for the enzymes of the molybdopterin biosynthetic pathway (*moaE*, *mobB*, and *moeB*) were upregulated. They might be involved in supplying the putative nitrate reductase with molybdopterin, while the formylmethanofuran dehydrogenase enzyme used for methanogenesis harbors a tungstopterin (42). The pathway of molybdopterin biosynthesis has been extensively studied, and it is highly conserved (63). The pathway of tungstopterin biosynthesis is thought to be homologous up to the step of metal insertion (64). It has been proposed that the biosynthetic machinery is able to distinguish between the two metals and insert the correct metal into the respective enzyme by employing MoeA isoenzymes selective for either molybdate or tungstate (64, 65). All archaeal genomes sequenced so far encode two *moeA* isoforms sharing around 40% identity (64). This might explain how organisms are able to correctly express different proteins with molybdopterin and tungstopterin simultaneously (64). However, this hypothesis still lacks experimental validation.

As shown in Fig. 4C and D, the expression of genes partaking in N$_2$ fixation and molybdenum acquisition remained unchanged at 21 h and 25 h; therefore, no differential expression was observed. The overall expression of genes in terms of TPM followed a similar pattern. An established marker for metabolic activity of methanogens, the transcription level of *mcrA* (66) was used as a reference point for comparison (Fig. 4B to D).

None of the methanogenesis pathway genes showed any difference in transcription levels at the early exponential phase. However, after 25 h nearly all genes involved in methanogenesis were downregulated. The expression of F$_{420}$-reducing hydrogenase was maintained at the same level at all time points to supply reduced F$_{420}$ from H$_2$ oxidation. Hydrogenotrophic *Methanococcales* can alternatively use formate as electron source (67, 68), and in this case a putative formate transporter (*fdhC*) and formate dehydrogenase (*fdhF*) were found to be downregulated at 3 h. In addition to methanogenesis, numerous anabolic processes were shut down at 25 h, such as carbon assimilation (e.g., pyruvate:ferredoxin oxidoreductase), amino acid metabolism (e.g., ketol-acid reductoisomerase), lipid biosynthesis (e.g., hydroxymethylglutaryl-coenzyme A synthase), ATP synthesis (i.e., ATP synthase), and vitamin and coenzyme biosynthesis (e.g., *hemE*). Such a decrease in catabolic and anabolic processes combined with the downregulation of genes involved in S-layer formation and cellular division (e.g., *ftsZ*) suggests a fine-tuned metabolic mode of energy saving to prioritize nitrogen fixation.

Transcriptional and translational machineries were considerably downregulated. Both the transcriptional and translational machineries responded negatively to diazotrophic conditions after 3 h. This included downregulation of RNA polymerase subunits (*rpoA$_2$H*), the sigma factor 70 (*rpoD*, controlling the transcription of housekeeping genes), and 35 ribosomal proteins (Table S2). Our rRNA expression data have the limitation that rRNA removal treatment was performed before sequencing. Still, we could detect downregulation of rRNA expression, which was corroborated by downregulation of ribosomal proteins. Some genes with a putative function in tRNA and ribosome biogenesis and biosynthesis of nucleotide and amino acid precursors were also negatively affected, corroborating a deep impact on the overall translation process under N$_2$-fixing conditions. For example, two key enzymes (transketolase and transaldolase) of the nonoxidative branch of the pentose phosphate pathway (69) for synthesis of nucleotides and histidine (from ribose 5-phosphate), as well as aromatic amino acids (from erythrose 4-phosphate) precursors were downregulated (Table S2). This was another notable difference from *Methanosarcina*, in which the upregulation of genes involved in the synthesis of aromatic amino acids has been observed (55). Notably, tRNA^{Thr} and tRNA^{Ser} were upregulated, while tRNA^{Ala} and tRNA^{Pro} were downregulated at 3 h. Taken together, changes in the expression of all the mentioned genes contributed to the restriction of the entire translation process. The DNA replication system seemed not to be impacted, since we could not observe any changes in the transcript levels of DNA polymerase-encoding genes. This was in line with the results of Belay and coworkers, who showed that during diazotrophic growth the cellular protein content of *M. thermolithotrophicus* was significantly reduced (7).

High expression of CRISPR-Cas genes and putative viral genes. Components of the CRISPR-Cas virus defense system were also upregulated under diazotrophy at 3 h, including Csm1 to -5 of the Csm effector complex and CRISPR-associated proteins (Cas proteins) Cas4 to -6 (Table S2). This might be explained by presence of a prophage that is expressed when the cells are energy depleted or otherwise stressed. The Phaster server, a tool for finding prophages in bacterial genomes (70), was unsuccessful in detecting any complete prophages in the *M. thermolithotrophicus* genome. However, we identified a locus of 28 cooccurring ORFs that were highly expressed at all time points (Fig. 5). Twenty of these genes had no characterized homologs, while eight encoded putative virus-like, replication, and mobile genetic elements. We used models generated by Alphafold2 (71) and a membrane prediction tool to gain further information on these sequences and examined 18 confident models (Fig. 5). Five of these proteins were predicted to be secreted or embedded in the membrane, but structural homologs were scarce based on the predicted models (Table S3). The protein encoded by the third gene of the locus (Fig. 5) has already been structurally characterized in *Thermococcales* (72) and is believed to be a virus-like element. While *Methanococcales* share some of these genes,
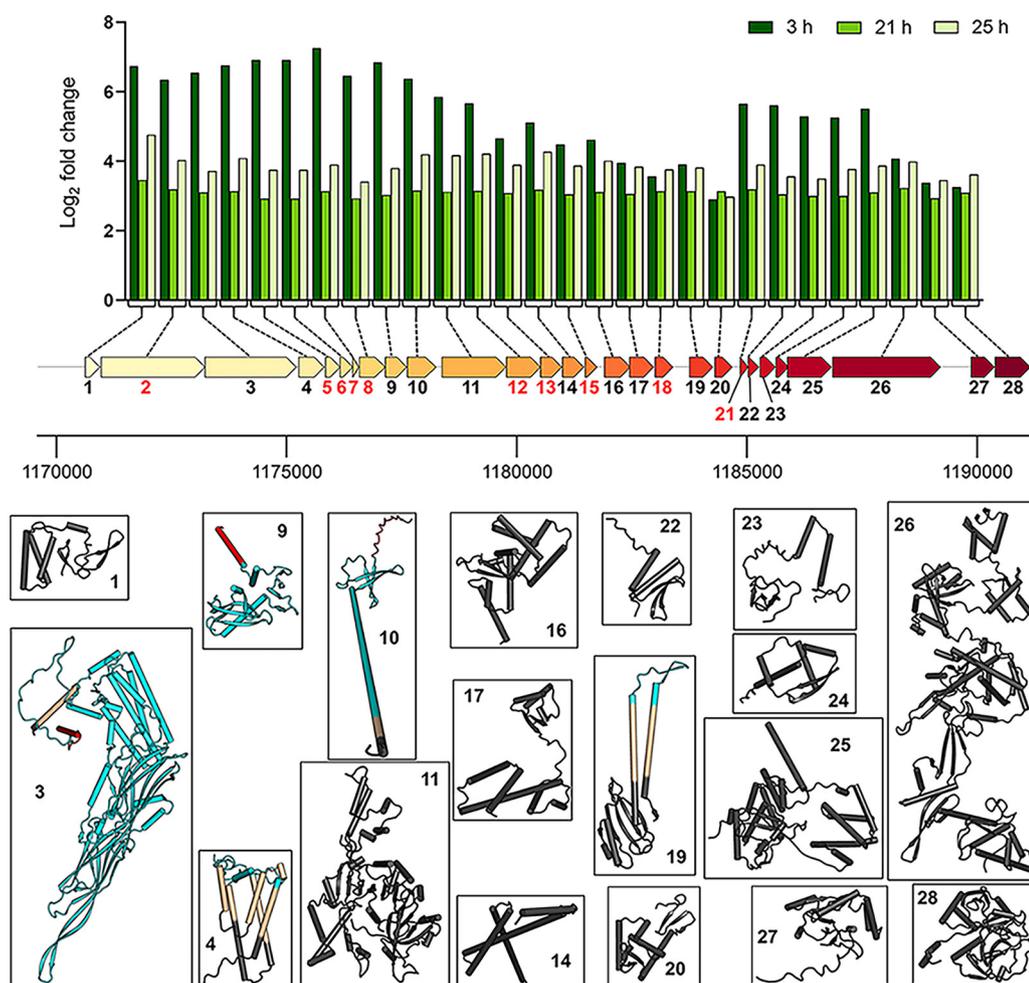
FIG 5 Log$_2$ fold changes over time for 28 cooccurring genes of unknown function that are of putative viral origin. AlphaFold 2 (71) models are represented by cartoons and color-coded as follows: black, intracellular segments; cyan, predicted extracellular segments; ocher, transmembrane segments; red, predicted signal sequences. The topology prediction was made via the DeepTMHMM server (109). Genes with numbered red labels highlight models with an overall per-residue confidence score (pLDDT) of <75. These models were not presented due to their low confidence scores.

Methanosarcina mazei, a model organism for the *Methanosarcinales*, contains only two of them in its genome (Table S3). These 28 cooccurring genes might be derived from a small plasmid transferred by conjugation or represent a yet-unknown prophage, although we could not identify a protein that could be responsible for independent insertion or replication of this element. This region could therefore be considered part of the dark matter in archaeal genomes (73). However, active transcription of all the genes of this locus might be a plausible explanation for the observed upregulation of the CRISPR-Cas system.

Another locus of 26 cooccurring ORFs was downregulated under diazotrophic conditions at all time points (Fig. S5). Again, most of these genes have unknown functions, with the exception of a putative transcriptional regulator, a mini-chromosome maintenance protein, and a recombinase (Table S4). Four of these genes were predicted to encode secreted proteins, and eight contained transmembrane segments. Based on these observations, we assume that these 26 cooccurring open reading frames are also derived from mobile elements or prophage-associated genes.

DISCUSSION

Diazotrophy allows microbes to survive when nitrogen becomes limiting in their natural habitats. This is also the case for (hyper)thermophilic methanogens, such as *M. thermolithotrophicus*, which has been shown to rely on diazotrophy in different environments (10–13).

A notable exception among *Methanococcales* is *Methanocaldococcus jannaschii*, a nondiazotrophic methanogen isolated from a white smoking crater on the East Pacific Rise. While ammonium concentrations at the *M. jannaschii* isolation site were not measured (74), some parts of the East Pacific Rise, such as Guaymas Basin, are known to feature notable ammonium concentrations (e.g., 15.3 mM [75]). Availability of this inorganic nitrogen source might have resulted in a complete loss of nitrogen-fixing abilities in *M. jannaschii*.

Due to its active N_2 fixation, *M. thermolithotrophicus* represents one of the contributors to the available nitrogen pool in specific environments (10–13). In our laboratory cultures, *M. thermolithotrophicus* released only minute amounts of up to 40 μM ammonia to the medium, which might have simply resulted from passive diffusion or cell lysis. It therefore seems that N_2 fixation *in M. thermolithotrophicus* is precisely controlled to avoid any losses, likely by the $NifI_{1,2}$ regulation system (Fig. 6). The excessive energy cost inflicted by nitrogenase activity was noticeable in our physiology experiments (Fig. 1), as it reduced final yields by a factor of 2 in both batch- and fermenter-grown cultures.

Transcriptome profiling under diazotrophic conditions revealed that *M. thermolithotrophicus* relies on multiple synergistic strategies that ensure both sufficient N_2 fixation and energy preservation to support cellular growth (Fig. 6). In addition, *M. thermolithotrophicus* enhances nitrogen acquisition by increasing ammonia uptake via its *amtB$_2$* transporter and glutamine synthetase overall activity, a strategy that has been described before based on proteomics of nitrogen-starved *M. maripaludis* cultures (59). While upregulation of the *nif* and *amtB/glnK* operons under nitrogen limitation has been previously reported in other diazotrophs (54, 55, 59), the upregulation of a putative nitrate transporter (*narK*), a nitrate reductase (*narB*), and a new isoform of F_{420}-dependent sulfite reductase (*fsr*) reported in this study is so far unique to *M. thermolithotrophicus*.

Like *M. thermolithotrophicus*, the hyperthermophile *Methanocaldococcus infernus* has also been reported to grow on nitrate as sole nitrogen source (76). *M. infernus* was isolated from a smoking crater of the Logatchev hydrothermal vent field (76, 77), from which other nitrate-reducing organisms have been isolated as well (78). Although hydrothermal fluids have been reported to be depleted in nitrate and nitrite (79), bottom seawater can contain nitrate for use as electron acceptor and nitrogen source. The enzyme which can possibly be used for intracellular nitrate reduction by these methanogens is the nitrate reductase NarB. NarB is expected to harbor a tungstopterin or molybdopterin cofactor, which might explain upregulation of molybdopterin biosynthesis genes (*moaE*, *mobB*, *moeB*) under diazotrophic conditions. In this context, the molybdate transporter Mod would play a dual role, supplying Mo for both the nitrogenase and NarB metallo-cofactor.

Strict dependence on Mo for diazotrophy has also been described for *Methanococcus maripaludis* (28). Likewise, our results support that the single nitrogenase operon of *M. thermolithotrophicus* encodes a molybdenum nitrogenase. This is corroborated by the *M. thermolithotrophicus* phylogenetic position (19), as well as high transcription levels of the *mod* operon (encoding a molybdate ABC transporter) adjacent to the nitrogenase operon under diazotrophy. In addition to *modA* within the *mod* operon, *M. thermolithotrophicus* features two additional *modA* instances. These are not colocalized with *modBC* genes and were not differentially expressed in our experiments. Although multiple *modA* genes are present in different diazotrophic organisms, including *A. vinelandii* and the methanogens *M. maripaludis* and *M. mazei*, the benefit of multiple *modA* genes is not yet understood. Interestingly, in *M. mazei* (55), no upregulation of Mod transporters was observed. In *Methanosarcinales*, expression of *modABC* is under the control of the transcriptional regulator ModE. This regulator also affects transcriptional regulation of different nitrogenases types, if these are present (80, 81). *Methanococcales* are devoid of *modE* homologs, which might explain the observed upregulation of the *mod* operon under diazotrophic growth in *M. thermolithotrophicus*.

The *nif* operon exhibited the highest log_2FC values between both conditions. However, changes in transcript levels of the *nifE* and *nifN* FeMo-co biosynthesis genes were less prominent than those of the structural *nif* genes. The former are possibly required in lower amounts, as was the case for the FeMo-co biosynthesis gene *nifB*. Levels of *nifX* transcription were similar to those of the *nifDK* nitrogenase genes, suggesting potential association with the nitrogenase
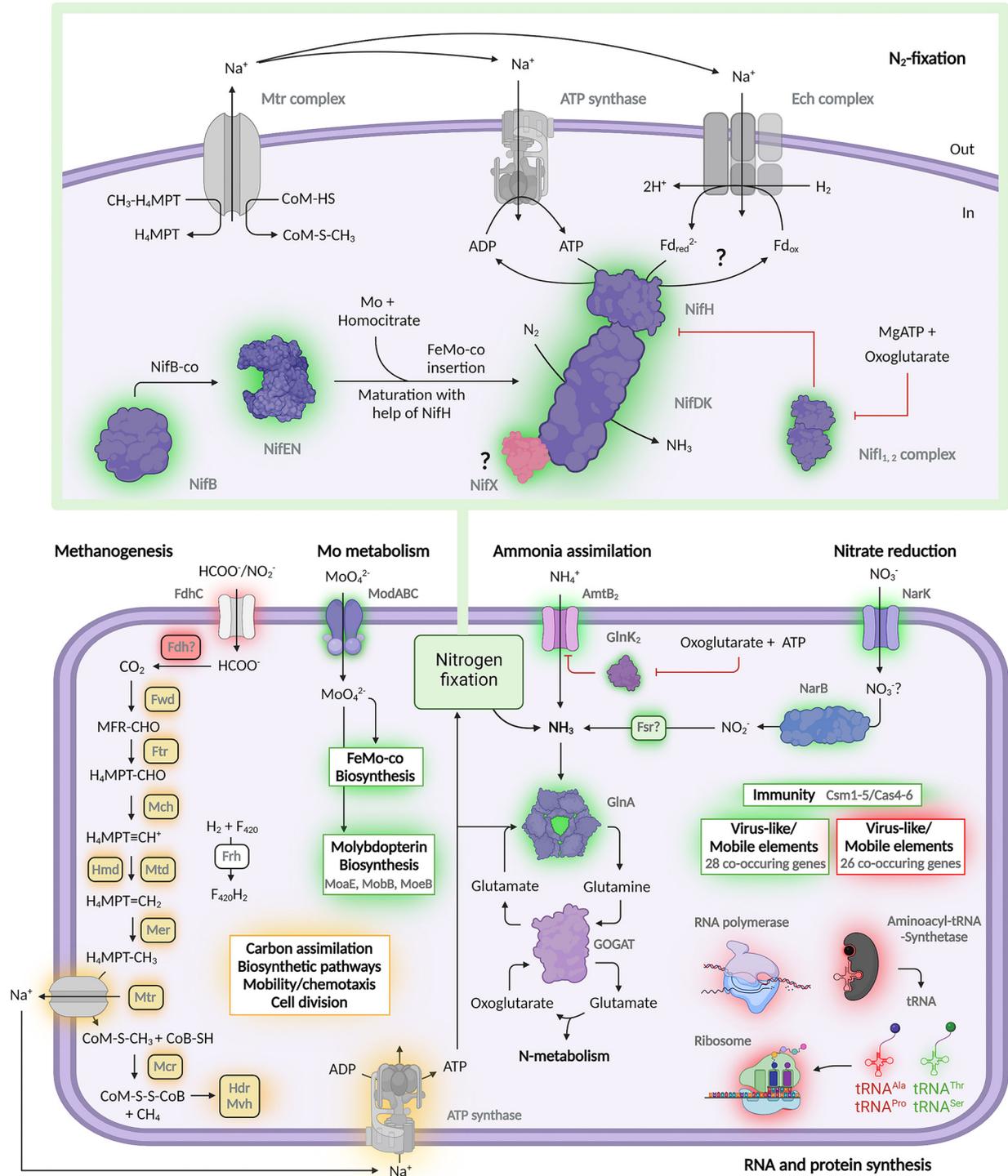
FIG 6 Influence of nitrogen fixation on the metabolism of *M. thermolithotrophicus*. In the scheme, differentially expressed genes involved in methanogenesis, nitrogen fixation, ammonia assimilation, nitrate reduction, molybdate import, molybdopterin biosynthesis, immunity, mobile elements, and transcription and translation, as well the tRNAs, are shown. Differentially upregulated and downregulated transcripts at 3 h postinoculation are highlighted with a green and red glow, respectively. An orange glow highlights differentially downregulated gene transcripts at 25 h postinoculation. Full names of enzymes can be found in Table S2. This picture was created with Biorender.

complex. The *nifX* gene is not essential for nitrogen fixation in *M. maripaludis* and has not been detected in the nitrogenase complex of this archaeon. Instead, it has been proposed to take on the role of the FeMo-co biosynthesis genes (17). However, since *M. maripaludis nifX* is not homologous to *M. thermolithotrophicus nifX*, future experimental evidence is required to clarify its putative function and association to NifDK in (hyper)thermophilic methanogens.

The enormous ATP investment required for diazotrophy did not affect the transcription of methanogenesis genes at the onset of diazotrophic growth. This observation corroborates previous studies (55, 59) that reported that the proportion of the overall energy required for maintenance increased when switching to diazotrophy (82). The high energy demand for nitrogen fixation is counterbalanced by drastic downregulation of the transcription and translation machineries. The question is whether this is an *a priori* and thus targeted metabolic adaptation, or a mere general consequence of energy starvation. It has been described that transcription and translation are reduced upon nutrient limitations. After all, 40 to 70% of the cellular ATP pool in growing bacteria is attributed to protein synthesis (83). It is also known that bacteria can balance tRNA abundances when under stress to selectively regulate the translation of stress-induced proteins. These proteins contribute to the response and adaptation to different types of stress, including nutrient limitation (84). It has recently been shown that *M. maripaludis* adopts a resource relocation strategy upon energy depletion (85): instead of reducing ribosome numbers, the cells rather redistribute available energy and decrease catabolic and ribosomal activities, a strategy also described in bacteria (86). Therefore, our results suggest a targeted response associated with switching to a diazotrophic lifestyle. In comparison, no such changes in the transcript levels of transcriptional and translational genes were detected in either *Azotobacter vinelandii* (54) or *M. mazei* (55) when switching to diazotrophic growth, even though in the case of the former 30% of the genes were differentially expressed (as observed by differential transcriptomics), and in the case of the latter, 5% were differentially expressed (as observed by microarray analysis).

Studies in *M. mazei* and *M. maripaludis* confirmed the prominent regulation of nitrogen metabolism at the transcriptional level, particularly the importance of the global nitrogen regulatory repressor NrpR of nitrogen assimilation genes (24, 25, 35, 87). NrpR represses transcription by binding to the *nif* and *glnA* promoter regions in a 2-oxoglutarate-dependent manner. NrpR from *M. thermolithotrophicus* shares 38.2% identity with the one from *M. mazei*. Interestingly, NrpR expression levels are not regulated by nitrogen supply (24, 25). Thus, although a homolog of NrpR does exist in *M. thermolithotrophicus*, its transcription level remained unchanged during diazotrophic growth. In contrast to NrpR, NrpA, the *nif* promoter-specific activator, sharing 32% identity with the one from *M. mazei*, is known to be upregulated upon nitrogen limitation (88). We did not observe this in our experiments, suggesting that NrpA could be constitutively expressed in *M. thermolithotrophicus*.

One major difference between *M. mazei* and *M. thermolithotrophicus*, however, is the role of regulatory sRNAs. *Methanosarcina mazei* Gö1 expresses multiple sRNA genes in response to nitrogen limitation (56), with sRNA$_{154}$ being the best characterized (36). In *M. mazei*, sRNA$_{154}$ stabilizes the nitrogenase, *glnA$_1$*, and *nrpA* mRNAs (36). While sRNA$_{154}$ is highly conserved within *Methanosarcinales*, there are no sRNA$_{154}$ homologs in *Methanococcales* (89). Consequently, we did not detect such a pattern in *M. thermolithotrophicus*.

Transcriptome analysis can provide valuable information on proteins involved in nitrogen fixation. For instance, transcriptome analysis allowed the discovery that in *A. vinelandii* electron bifurcation is coupled to a flavodoxin that fuels the nitrogenase (54). However, we could not detect any putative novel candidates that could shuttle electrons to the nitrogenase, and we propose reduced ferredoxin as a candidate for electron delivery, a known electron carrier for anabolic reactions. *M. thermolithotrophicus* would use H$_2$ oxidation for ferredoxin reduction, an endergonic process that requires the coupling of the influx of sodium ions by the Ech complex. The reduced ferredoxin would then drive the reduction of NifH for N$_2$ fixation (Fig. 6).

Here, we have presented novel insights into the metabolic rebalancing that methanogens employ to accommodate Mo-dependent diazotrophy. Our findings include additional cellular responses to this stress, such as the unexpected CRISPR-Cas upregulation. This upregulation is most probably caused by the expression of putative virus-like elements (Fig. 6). Further studies at the protein level are required to decipher the mechanism of this complex adaptation in greater detail. In particular, molecular investigations of the regulatory functions of NifI$_{1,2}$, the role of NifX, and the intrinsic properties of the thermostable NifHDK complex will unveil the secrets of the astonishing diazotrophic capabilities of *M. thermolithotrophicus*.

MATERIALS AND METHODS

Growth conditions. *Methanothermococcus thermolithotrophicus* DSM 2095 (Leibniz Institute DSMZ, Braunschweig, Germany) was grown under anoxic conditions in minimal mineral medium with 1×10^5 Pa overpressure of either H$_2$:CO$_2$ (80%:20%) for nondiazotrophic cultures or 1.2×10^5 Pa overpressure H$_2$:CO$_2$:N$_2$ (58.2%:14.5%:27.3%) for diazotrophic cultures. Cultures were grown in 250-mL serum flasks (Glasgerätebau Ochs, Bovenden, Germany) sealed with rubber stoppers and aluminum crimps in final volumes of 10 mL with a 1:10 inoculum. Serum flasks and media were made anoxic prior to inoculation by sparging with N$_2$ and two final gas exchanges with H$_2$:CO$_2$ (80%:20%). Incubation was done at 65°C, in the dark, without shaking.

Medium composition. The minimal mineral medium was prepared as described by Müller et al. (40), but with replacement at an equal final concentration of Fe(NH$_4$)$_2$(SO$_4$)$_2$·12H$_2$O by FeCl$_2$·4H$_2$O and of Na$_2$SeO$_3$·5H$_2$O by Na$_2$SeO$_4$. The used trace metal solution when 100-fold concentrated contained 7.1 mM nitrilo-triacetic acid, 0.45 mM MnCl$_2$·2H$_2$O, 0.68 mM FeCl$_3$·6H$_2$O, 0.41 mM CaCl$_2$, 0.76 mM CoCl$_2$, 0.66 mM ZnSO$_4$·6H$_2$O, 0.28 mM CuSO$_4$, 0.19 mM Na$_2$MoO$_4$·2H$_2$O, 0.38 mM NiCl$_2$·6H$_2$O and 0.19 mM VCl$_3$. The final pH was adjusted to 6.0 by addition of NaOH pellets. The final media were subsequently made anoxic by several degassing and N$_2$ addition cycles (minimum, 25 cycles). The same medium, but without NH$_4$Cl, was used for diazotrophic cultures. Na$_2$S was used as both a reductant and sulfur source at a final concentration of 1.5 mM in all cases.

Adaptation to diazotrophic conditions. *M. thermolithotrophicus* was adapted to growth under diaz-otrophic conditions after NH$_4$Cl depletion from the media by three successive transfers to the same media without NH$_4$Cl, with a headspace containing 1.2×10^5 Pa H$_2$:CO$_2$:N$_2$ (58.2%:14.5%:27.3%) as described above. Cultures grown in media without NH$_4$Cl with a gas phase of 1×10^5 Pa H$_2$:CO$_2$ (80%:20%) were used as nega-tive controls.

Influence of trace metal availability on diazotrophic growth and tungstate inhibition. To deter-mine influences of Mo and V on diazotrophic growth, we depleted media of Mo and V by three successive transfers of diazotrophic *M. thermolithotrophicus* cultures to media prepared as described above, but without Na$_2$MoO$_4$·2H$_2$O or VCl$_3$ or without both. The minimal concentration of Mo for diazotrophic growth was deter-mined in a series of incubations, in which an already Mo-depleted diazotrophic culture was supplemented with Na$_2$MoO$_4$·2H$_2$O concentrations ranging from 0.01 to 100 μM. W inhibition was tested by supplementing the already W-depleted diazotrophic culture with 0.001 to 10 μM Na$_2$WO$_4$·2H$_2$O; a supplementation with the concentration of 100 μM gave unreproducible results (data not shown). W depletion was done as described above, with transfers to the media prepared without Na$_2$WO$_4$·2H$_2$O.

Cultivation in a fermenter. *M. thermolithotrophicus* was continuously grown in a 10-liter fermenter (BIOSTAT B plus; Sartorius, Göttingen, Germany) under diazotrophic conditions with either 10 mM Na$_2$SO$_3$ or Na$_2$SO$_4$ as sulfur source instead of Na$_2$S. The final culture volumes were 7 liters for the culture grown with Na$_2$SO$_3$ and 6 liters for the culture grown with Na$_2$SO$_4$. The culture was continuously sparged with H$_2$:CO$_2$ (80%:20%) and N$_2$ in the ratio of 1:1 and stirred with the speed of 500 rpm, at 65°C. As an inoculum, cultures cultivated in the same media were used in a ratio of 1:10.

Ammonia measurement. Ammonia concentrations were measured in the culture supernatant. Aliquots of 0.5 mL of culture were subsampled aerobically at each time point, cells were pelleted by cen-trifugation for 5 min at 15,700 $\times$ *g* using 5415R microcentrifuge (Eppendorf, Hamburg, Germany), and the supernatant was frozen at $-20°$ C until further use. Ammonia was measured by the salicylate-nitro-prussidine method (90) in ROTILABO F-profile microtitration plates (Carl Roth GmbH, Karlsruhe, Germany). Standards in the range from 0 to 600 μM NH$_4$Cl were prepared in the same medium used for cultivation. An 80 μL volume of salicylate reagent (424.7 mM sodium salicylate, 193.8 mM trisodium ci-trate dihydrate, 193.8 mM disodium tartrate dihydrate, and 0.95 mM sodium nitroprusside dihydrate) and 80 μL volume hypochlorite reagent (10% sodium hypochlorite and 1.5 M NaOH mixed in the ratio of 1:36) were added to 40 μL of each sample. The plate was additionally mixed for 5 min on a shaker and incubated in the dark at the room temperature for 45 min. The absorbance was measured at 650 nm using an Infinite 200 PRO plate reader (Tecan, Männedorf, Switzerland) at room temperature.

Evolutionary analyses. NifHDK, VnfHDK, and AnfHDK sequences of 35 selected species were aligned using MUSCLE (91) (default parameters) in MEGA11 (92, 93). Afterwards, ambiguous positions were removed for each possible pairing (pairwise deletion option). The final alignment contained a total of 485 NifH positions, 729 NifD positions, and 563 NifK positions. The sequence's evolutionary history was inferred using the neigh-bor-joining method (94) with the JTT matrix model for multiple substitutions (95). The analysis was conducted in MEGA11 with ChlLNB from *Chlorobium limicola* as outgroup.

Transcriptomics setup. The cultures for transcriptomic profiling were grown in triplicates in batches as described above, but scaled up to culture volumes of 60 mL in 1-liter pressure-resistant Duran bottles. Inocula used to start the cultures were adapted to the respective conditions prior to inoculation, as described above. Subsamples were taken anaerobically on ice after 3 h, 21 h, and 25 h and transferred to an anaerobic chamber, where they were pelleted and subsequently frozen in liquid nitrogen immedi-ately after being taken out of the anaerobic chamber. Samples were stored at $-80°$C until further use. Samples were shipped on dry ice to the Max Planck Genome Centre in Cologne, where they were proc-essed and sequenced. RNA extraction and quality control, including rRNA removal, were also done at the Max Planck Genome Centre in Cologne.

Transcriptome sequencing and analysis. Transcriptome sequencing was performed on an Illumina HiSeq 3000 system (San Diego, CA, USA). Information on the raw reads is summarized in Table S1. Raw RNA reads were quality trimmed and repaired using the bbduk and repair.sh scripts of the BBMap v35.14 suite (https://sourceforge.net/projects/bbmap/). Reads with a minimum length of 70 bp and a quality score of 20 were filtered for rRNAs using SortMeRNA v3.0.3 (96). Remaining mRNA reads were mapped against the *M. ther-molithotrophicus* reference genome using Bowtie2 (97) as part of the SqueezeMeta v1.3.1 pipeline (98). The

DESeq2 R package (99) was subsequently used to calculate log$_2$ fold changes, standard errors, test statistics, and adjusted *P* values. Changes in expression levels with adjusted *P* values (P_{adj}) of <0.05 and a minimum 2-fold change ratio (log$_2$ fold change of ≥1) were considered significant (54).

The high-quality genome of *M. thermolithotrophicus* was used as a reference for mapping RNA reads. The genomic DNA of *M. thermolithotrophicus* was isolated using the protocol from Platero et al. (100) and sequenced on PacBio Sequel II platform using a single SMRT cell at the Max Planck Genome Center in Cologne. The genome was assembled using Flye v2.7 (101) and annotated as part of the SqueezeMeta pipeline. Briefly, the ORFs were predicted using Prodigal (102), and similarity searches for GenBank (103), eggNOG (104), and KEGG (105), were done using Diamond (106). HMM homology searches were done with HMMER3 (107) for the Pfam database (108).

Structural modeling and bioinformatic analysis. Alphafold2 was run with default parameters for all generated models. Predictions of membrane regions and overall topology were run on the DeepTMHMM server (https://dtu.biolib.com/DeepTMHMM/) (109) as of 20 June 2022.

Data availability. The *M. thermolithotrophicus* genome sequence and the transcriptome raw reads are available under ENA project number PRJEB53446.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.
FIG S1, TIF file, 0.7 MB.
FIG S2, TIF file, 1.6 MB.
FIG S3, TIF file, 2.5 MB.
FIG S4, TIF file, 0.1 MB.
FIG S5, TIF file, 1.1 MB.
TABLE S1, XLSX file, 0.01 MB.
TABLE S2, XLSX file, 0.1 MB.
TABLE S3, XLSX file, 0.01 MB.
TABLE S4, XLSX file, 0.01 MB.
TABLE S5, XLSX file, 0.01 MB.

REFERENCES

1. Thauer RK, Kaster AK, Seedorf H, Buckel W, Hedderich R. 2008. Methanogenic archaea: ecologically relevant differences in energy conservation. Nat Rev Microbiol 6:579–591. https://doi.org/10.1038/nrmicro1931.

2. Shima S, Huang G, Wagner T, Ermler U. 2020. Structural basis of hydrogenotrophic methanogenesis. Annu Rev Microbiol 74:713–733. https://doi.org/10.1146/annurev-micro-011720-122807.

3. Offre P, Spang A, Schleper C. 2013. Archaea in biogeochemical cycles. Annu Rev Microbiol 67:437–457. https://doi.org/10.1146/annurev-micro-092412-155614.

4. Conrad R. 2020. Importance of hydrogenotrophic, aceticlastic and methylotrophic methanogenesis for methane production in terrestrial, aquatic and other anoxic environments: a mini review. Pedosphere 30:25–39. https://doi.org/10.1016/S1002-0160(18)60052-9.

5. Buckel W, Thauer RK. 2013. Energy conservation via electron bifurcating ferredoxin reduction and proton/Na(+) translocating ferredoxin oxidation. Biochim Biophys Acta 1827:94–113. https://doi.org/10.1016/j.bbabio.2012.07.002.

6. Lemaire ON, Jespersen M, Wagner T. 2020. CO$_2$-fixation strategies in energy extremophiles: what can we learn from acetogens? Front Microbiol 11:486. https://doi.org/10.3389/fmicb.2020.00486.

7. Belay N, Sparling R, Daniels L. 1984. Dinitrogen fixation by a thermophilic methanogenic bacterium. Nature 312:286–288. https://doi.org/10.1038/312286a0.

8. Murray PA, Zinder SH. 1984. Nitrogen fixation by a methanogenic archaebacterium. Nature 312:284–286. https://doi.org/10.1038/312284a0.

9. Harris DF, Lukoyanov DA, Kallas H, Trncik C, Yang Z-Y, Compton P, Kelleher N, Einsle O, Dean DR, Hoffman BM, Seefeldt LC. 2019. Mo-, V-, and Fe-nitrogenases use a universal eight-electron reductive-elimination mechanism to achieve N$_2$ reduction. Biochemistry 58:3293–3301. https://doi.org/10.1021/acs.biochem.9b00468.

10. Mehta MP, Butterfield DA, Baross JA. 2003. Phylogenetic diversity of nitrogenase (*nifH*) genes in deep-sea and hydrothermal vent environments of the Juan de Fuca Ridge. Appl Environ Microbiol 69:960–970. https://doi.org/10.1128/AEM.69.2.960-970.2003.

11. Man-Aharonovich D, Kress N, Zeev EB, Berman-Frank I, Béjà O. 2007. Molecular ecology of *nifH* genes and transcripts in the eastern Mediterranean Sea. Environ Microbiol 9:2354–2363. https://doi.org/10.1111/j.1462-2920.2007.01353.x.

12. Dang H, Luan X, Zhao J, Li J. 2009. Diverse and novel *nifH* and *nifH*-like gene sequences in the deep-sea methane seep sediments of the Okhotsk Sea. Appl Environ Microbiol 75:2238–2245. https://doi.org/10.1128/AEM.02556-08.

13. Bae HS, Morrison E, Chanton JP, Ogram A. 2018. Methanogens are major contributors to nitrogen fixation in soils of the Florida Everglades. Appl Environ Microbiol 84:e02222-17. https://doi.org/10.1128/AEM.02222-17.

14. Dekas AE, Poretsky RS, Orphan VJ. 2009. Deep-sea archaea fix and share nitrogen in methane-consuming microbial consortia. Science 326:422–426. https://doi.org/10.1126/science.1178223.

15. Mus F, Alleman AB, Pence N, Seefeldt LC, Peters JW. 2018. Exploring the alternatives of biological nitrogen fixation. Metallomics 10:523–538. https://doi.org/10.1039/c8mt00038g.

16. Kessler PS, Blank C, Leigh JA. 1998. The *nif* gene operon of the methanogenic archaeon *Methanococcus maripaludis*. J Bacteriol 180:1504–1511. https://doi.org/10.1128/JB.180.6.1504-1511.1998.

17. Kessler PS, Leigh JA. 1999. Genetics of nitrogen regulation in *Methanococcus maripaludis*. Genetics 152:1343–1351. https://doi.org/10.1093/genetics/152.4.1343.

18. Halbleib CM, Ludden PW. 2000. Regulation of biological nitrogen fixation. J Nutr 130:1081–1084. https://doi.org/10.1093/jn/130.5.1081.

19. Boyd ES, Hamilton TL, Peters JW. 2011. An alternative path for the evolution of biological nitrogen fixation. Front Microbiol 2:205. https://doi.org/10.3389/fmicb.2011.00205.

20. Boyd ES, Anbar AD, Miller S, Hamilton TL, Lavin M, Peters JW. 2011. A late methanogen origin for molybdenum-dependent nitrogenase. Geobiology 9:221–232. https://doi.org/10.1111/j.1472-4669.2011.00278.x.

21. Boyd ES, Peters JW. 2013. New insights into the evolutionary history of biological nitrogen fixation. Front Microbiol 4:201–212.

22. Boyd ES, Garcia Costas AM, Hamilton TL, Mus F, Peters JW. 2015. Evolution of molybdenum nitrogenase during the transition from anaerobic to aerobic metabolism. J Bacteriol 197:1690–1699. https://doi.org/10.1128/JB.02611-14.

23. Leigh JA, Dodsworth JA. 2007. Nitrogen regulation in bacteria and archaea. Annu Rev Microbiol 61:349–377. https://doi.org/10.1146/annurev.micro.61.080706.093409.

24. Lie TJ, Leigh JA. 2003. A novel repressor of *nif* and *glnA* expression in the methanogenic archaeon *Methanococcus maripaludis*. Mol Microbiol 47:235–246. https://doi.org/10.1046/j.1365-2958.2003.03293.x.

25. Lie TJ, Wood GE, Leigh JA. 2005. Regulation of *nif* expression in *Methanococcus maripaludis*: roles of the euryarchaeal repressor NrpR, 2-oxoglutarate, and two operators. J Biol Chem 280:5236–5241. https://doi.org/10.1074/jbc.M411778200.

26. Kessler PS, Daniel C, Leigh JA. 2001. Ammonia switch-off of nitrogen fixation in the methanogenic archaeon *Methanococcus maripaludis*: mechanistic features and requirement for the novel GlnB homologues, NifI$_1$ and NifI$_2$. J Bacteriol 183:882–889. https://doi.org/10.1128/JB.183.3.882-889.2001.

27. Blank CE, Kessler PS, Leigh JA. 1995. Genetics in methanogens: transposon insertion mutagenesis of a *Methanococcus maripaludis nifH* gene. J Bacteriol 177:5773–5777. https://doi.org/10.1128/jb.177.20.5773-5777.1995.

28. Kessler PS, McLarnan J, Leigh JA. 1997. Nitrogenase phylogeny and the molybdenum dependence of nitrogen fixation in *Methanococcus maripaludis*. J Bacteriol 179:541–543. https://doi.org/10.1128/jb.179.2.541-543.1997.

29. Dodsworth JA, Cady NC, Leigh JA. 2005. 2-Oxoglutarate and the PII homologues NifI$_1$ and NifI$_2$ regulate nitrogenase activity in cell extracts of *Methanococcus maripaludis*. Mol Microbiol 56:1527–1538. https://doi.org/10.1111/j.1365-2958.2005.04621.x.

30. Dodsworth JA, Leigh JA. 2006. Regulation of nitrogenase by 2-oxoglutarate-reversible, direct binding of a PII-like nitrogen sensor protein to dinitrogenase. Proc Natl Acad Sci U S A 103:9779–9784. https://doi.org/10.1073/pnas.0602278103.

31. Lobo AL, Zinder SH. 1990. Nitrogenase in the archaebacterium *Methanosarcina barkeri* 227. J Bacteriol 172:6789–6796. https://doi.org/10.1128/jb.172.12.6789-6796.1990.

32. Lobo AL, Zinder SH. 1988. Diazotrophy and nitrogenase activity in the archaebacterium *Methanosarcina barkeri* 227. Appl Environ Microbiol 54:1656–1661. https://doi.org/10.1128/aem.54.7.1656-1661.1988.

33. Chien Y-T, Zinder SH. 1996. Cloning, functional organization, transcript studies, and phylogenetic analysis of the complete nitrogenase structural genes (*nifHDK2*) and associated genes in the archaeon *Methanosarcina barkeri* 227. J Bacteriol 178:143–148. https://doi.org/10.1128/jb.178.1.143-148.1996.

34. Ehlers C, Veit K, Gottschalk G, Schmitz RA. 2002. Functional organization of a single *nif* cluster in the mesophilic archaeon *Methanosarcina mazei* strain Gö1. Archaea 1:143–150. https://doi.org/10.1155/2002/362813.

35. Weidenbach K, Ehlers C, Kock J, Ehrenreich A, Schmitz RA. 2008. Insights into the NrpR regulon in *Methanosarcina mazei* Gö1. Arch Microbiol 190:319–332. https://doi.org/10.1007/s00203-008-0369-3.

36. Prasse D, Förstner KU, Jäger D, Backofen R, Schmitz RA. 2017. sRNA$_{154}$ a newly identified regulator of nitrogen fixation in *Methanosarcina mazei* strain Gö1. RNA Biol 14:1544–1558. https://doi.org/10.1080/15476286.2017.1306170.

37. Mehta MP, Baross JA. 2006. Nitrogen fixation at 92°C by a hydrothermal vent archaeon. Science 314:1783–1786. https://doi.org/10.1126/science.1134772.

38. Huber H, Thomm M, König H, Thies G, Stetter KO. 1982. *Methanococcus thermolithotrophicus*, a novel thermophilic lithotrophic methanogen. Arch Microbiol 132:47–50. https://doi.org/10.1007/BF00690816.

39. Jespersen M, Pierik AJ, Wagner T. 2022. The structure of the F$_{420}$-dependent sulfite-detoxifying enzyme from *Methanococcales* reveals a prototypical sulfite-reductase with assimilatory traits. bioRxiv. https://doi.org/10.1101/2022.04.07.487323.

40. Müller MC, Wagner T. 2021. The oxoglutarate binding site and regulatory mechanism are conserved in ammonium transporter inhibitors GlnKs from *Methanococcales*. Int J Mol Sci 22:8631. https://doi.org/10.3390/ijms22168631.

41. Balderston WL, Payne WJ. 1976. Inhibition of methanogenesis in salt marsh sediments and whole-cell suspensions of methanogenic bacteria by nitrogen oxides. Appl Environ Microbiol 32:264–269. https://doi.org/10.1128/aem.32.2.264-269.1976.

42. Wagner T, Ermler U, Shima S. 2018. Formyl-methanofuran dehydrogenase, p 1–18. *In* Scott RA (ed), Encyclopedia of inorganic and bioinorganic chemistry. John Wiley & Sons, Ltd., Hoboken, NJ.

43. Benemann JR, Smith GM, Kostel PJ, McKenna CE. 1973. Tungsten incorporation into *Azotobacter vinelandii* nitrogenase. FEBS Lett 29:219–221. https://doi.org/10.1016/0014-5793(73)80023-7.

44. Bulen WA. 1961. Effect of tungstate on the uptake and function of molybdate in *Azotobacter agilis*. J Bacteriol 82:130–134. https://doi.org/10.1128/jb.82.1.130-134.1961.

45. Nagatani HH, Brill WJ. 1974. Nitrogenase v. the effect of Mo, W and V on the synthesis of nitrogenase components in *Azotobacter vinelandii*. Biochim Biophys Acta 362:160–166. https://doi.org/10.1016/0304-4165(74)90037-3.

46. Siemann S, Schneider K, Oley M, Müller A. 2003. Characterization of a tungsten-substituted nitrogenase isolated from *Rhodobacter capsulatus*. Biochemistry 42:3846–3857. https://doi.org/10.1021/bi0270790.

47. Grunden AM, Shanmugam KT. 1997. Molybdate transport and regulation in bacteria. Arch Microbiol 168:345–354. https://doi.org/10.1007/s002030050508.

48. Self WT, Grunden AM, Hasona A, Shanmugam KT. 2001. Molybdate transport. Res Microbiol 152:311–321. https://doi.org/10.1016/s0923-2508(01)01202-5.

49. Aguilar-Barajas E, Díaz-Pérez C, Ramírez-Díaz MI, Riveros-Rosas H, Cervantes C. 2011. Bacterial transport of sulfate, molybdate, and related oxyanions. Biometals 24:687–707. https://doi.org/10.1007/s10534-011-9421-x.

50. Makdessi K, Andreesen JR, Pich A. 2001. Tungstate uptake by a highly specific ABC transporter in *Eubacterium acidaminophilum*. J Biol Chem 276:24557–24564. https://doi.org/10.1074/jbc.M101293200.

51. Bevers LE, Hagedoorn PL, Krijger GC, Hagen WR. 2006. Tungsten transport protein A (WtpA) in *Pyrococcus furiosus*: the first member of a new class of tungstate and molybdate transporters. J Bacteriol 188:6498–6505. https://doi.org/10.1128/JB.00548-06.

52. Balan A, Santacruz-Pérez C, Moutran A, Ferreira LC, Neshich G, Gonçalves Barbosa JA. 2008. Crystallographic structure and substrate-binding interactions of the molybdate-binding protein of the phytopathogen *Xanthomonas axonopodis* pv. *citri*. Biochim Biophys Acta 1784:393–399. https://doi.org/10.1016/j.bbapap.2007.11.013.

53. Imperial J, Hadi M, Amy NK. 1998. Molybdate binding by ModA, the periplasmic component of the *Escherichia coli mod* molybdate transport system. Biochim Biophys Acta 1370:337–346. https://doi.org/10.1016/s0005-2736(98)00003-0.

54. Hamilton TL, Ludwig M, Dixon R, Boyd ES, Dos Santos PC, Setubal JC, Bryant DA, Dean DR, Peters JW. 2011. Transcriptional profiling of nitrogen fixation in *Azotobacter vinelandii*. J Bacteriol 193:4477–4486. https://doi.org/10.1128/JB.05099-11.

55. Veit K, Ehlers C, Ehrenreich A, Salmon K, Hovey R, Gunsalus RP, Deppenmeier U, Schmitz RA. 2006. Global transcriptional analysis of *Methanosarcina mazei* strain Gö1 under different nitrogen availabilities. Mol Gen Genomics 276:41–55. https://doi.org/10.1007/s00438-006-0117-9.

56. Jäger D, Sharma CM, Thomsen J, Ehlers C, Vogel J, Schmitz RA. 2009. Deep sequencing analysis of the *Methanosarcina mazei* Gö1 transcriptome in response to nitrogen availability. Proc Natl Acad Sci U S A 106:21878–21882. https://doi.org/10.1073/pnas.0909051106.

57. Jenner LP, Cherrier MV, Amara P, Rubio LM, Nicolet Y. 2021. An unexpected P-cluster like intermediate *en route* to the nitrogenase FeMo-co. Chem Sci 12:5269–5274. https://doi.org/10.1039/d1sc00289a.

58. Kang W, Rettberg LA, Stiebritz MT, Jasniewski AJ, Tanifuji K, Lee CC, Ribbe MW, Hu Y. 2021. X-ray crystallographic analysis of NifB with a full complement of clusters: structural insights into the radical SAM-dependent carbide insertion during nitrogenase cofactor assembly. Angew Chem Int Ed Engl 60:2364–2370. https://doi.org/10.1002/anie.202011367.

59. Xia Q, Wang T, Hendrickson EL, Lie TJ, Hackett M, Leigh JA. 2009. Quantitative proteomics of nutrient limitation in the hydrogenotrophic methanogen *Methanococcus maripaludis*. BMC Microbiol 9:149. https://doi.org/10.1186/1471-2180-9-149.

60. Moreno-Vivián C, Cabello P, Martínez-Luque M, Blasco R, Castillo F. 1999. Prokaryotic nitrate reduction: molecular properties and functional distinction among bacterial nitrate reductases. J Bacteriol 181:6573–6584. https://doi.org/10.1128/JB.181.21.6573-6584.1999.

61. Duin EC, Wagner T, Shima S, Prakash D, Cronin B, Yanez-Ruiz DR, Duval S, Rumbeli R, Stemmler RT, Thauer RK, Kindermann M. 2016. Mode of action uncovered for the specific reduction of methane emissions from ruminants by the small molecule 3-nitrooxypropanol. Proc Natl Acad Sci U S A 113:6172–6177. https://doi.org/10.1073/pnas.1600298113.

62. Belay N, Jung KY, Rajagopal BS, Kremer JD, Daniels L. 1990. Nitrate as a sole nitrogen source for *Methanococcus thermolithotrophicus* and its effect on growth of several methanogenic bacteria. Curr Microbiol 21:193–198. https://doi.org/10.1007/BF02092121.

63. Schwarz G. 2005. Molybdenum cofactor biosynthesis and deficiency. Cell Mol Life Sci 62:2792–2810. https://doi.org/10.1007/s00018-005-5269-y.

64. Bevers LE, Hagedoorn P-L, Hagen WR. 2009. The bioinorganic chemistry of tungsten. Coord Chem Rev 253:269–290. https://doi.org/10.1016/j.ccr.2008.01.017.

65. Seelmann CS, Willistein M, Heider J, Boll M. 2020. Tungstoenzymes: occurrence, catalytic diversity and cofactor synthesis. Inorganics 8:44. https://doi.org/10.3390/inorganics8080044.

66. Alvarado A, Montañez-Hernández LE, Palacio-Molina SL, Oropeza-Navarro R, Luévanos-Escareño MP, Balagurusamy N. 2014. Microbial trophic interactions and *mcrA* gene expression in monitoring of anaerobic digesters. Front Microbiol 5:597. https://doi.org/10.3389/fmicb.2014.00597.

67. Lupa B, Hendrickson EL, Leigh JA, Whitman WB. 2008. Formate-dependent H₂ production by the mesophilic methanogen *Methanococcus maripaludis*. Appl Environ Microbiol 74:6584–6590. https://doi.org/10.1128/AEM.01455-08.

68. Wood GE, Haydock AK, Leigh JA. 2003. Function and regulation of the formate dehydrogenase genes of the methanogenic archaeon *Methanococcus maripaludis*. J Bacteriol 185:2548–2554. https://doi.org/10.1128/JB.185.8.2548-2554.2003.

69. 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.

70. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44:W16–W21. https://doi.org/10.1093/nar/gkw387.

71. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596:583–589. https://doi.org/10.1038/s41586-021-03819-2.

72. Keller J, Leulliot N, Soler N, Collinet B, Vincentelli R, Forterre P, van Tilbeurgh H. 2009. A protein encoded by a new family of mobile elements from Euryarchaea exhibits three domains with novel folds. Protein Sci 18:850–855. https://doi.org/10.1002/pro.73.

73. Makarova KS, Wolf YI, Forterre P, Prangishvili D, Krupovic M, Koonin EV. 2014. Dark matter in archaeal genomes: a rich source of novel mobile elements, defense systems and secretory complexes. Extremophiles 18:877–893. https://doi.org/10.1007/s00792-014-0672-7.

74. Jones WJ, Leigh JA, Mayer F, Woese CR, Wolfe RS. 1983. *Methanococcus jannaschii* sp. nov., an extremely thermophilic methanogen from a submarine hydrothermal vent. Arch Microbiol 136:254–261. https://doi.org/10.1007/BF00425213.

75. Wankel SD, Bourbonnais A, Charoenpong C. 2017. Microbial nitrogen cycling processes at submarine hydrothermal vents, p 179–222. *In* Jens K (ed), Life at Vents and Seeps, vol 2. De Gruyter, Berlin, Germany.

76. Jeanthon C, L'Haridon S, Reysenbach AL, Vernet M, Messner P, Sleytr UB, Prieur D. 1998. *Methanococcus infernus* sp. nov., a novel hyperthermophilic lithotrophic methanogen isolated from a deep-sea hydrothermal vent. Int J Syst Bacteriol 48:913–919. https://doi.org/10.1099/00207713-48-3-913.

77. Petersen S, Kuhn K, Kuhn T, Augustin N, Hékinian R, Franz L, Borowski C. 2009. The geological setting of the ultramafic-hosted Logatchev hydrothermal field (14°45′N, Mid-Atlantic Ridge) and its influence on massive sulfide formation. Lithos 112:40–56. https://doi.org/10.1016/j.lithos.2009.02.008.

78. Miroshnichenko ML, Kostrikina NA, Chernyh NA, Pimenov NV, Tourova TP, Antipov AN, Spring S, Stackebrandt E, Bonch-Osmolovskaya EA. 2003. *Caldithrix abyssi* gen. nov., sp. nov., a nitrate-reducing, thermophilic, anaerobic bacterium isolated from a Mid-Atlantic Ridge hydrothermal vent, represents a novel bacterial lineage. Int J Syst Evol Microbiol 53:323–329. https://doi.org/10.1099/ijs.0.02390-0.

79. Vetriani C, Voordeckers JW, Crespo-Medina M, O'Brien CE, Giovannelli D, Lutz RA. 2014. Deep-sea hydrothermal vent *Epsilonproteobacteria* encode a conserved and widespread nitrate reduction pathway (Nap). ISME J 8:1510–1521. https://doi.org/10.1038/ismej.2013.246.

80. Studholme DJ, Pau RN. 2003. A DNA element recognised by the molybdenum-responsive transcription factor ModE is conserved in Proteobacteria, green sulphur bacteria and Archaea. BMC Microbiol 3:24. https://doi.org/10.1186/1471-2180-3-24.

81. Chanderban M, Hill CA, Dhamad AE, Lessner DJ. 2021. *Methanosarcina acetivorans* simultaneously produces molybdenum, vanadium, and iron-only nitrogenases in response to fixed nitrogen and molybdenum depletion. bioRxiv. https://doi.org/10.1101/2021.06.03.447018.

82. Fardeau M-L, Peillex J-P, Belaïch J-P. 1987. Energetics of the growth of *Methanobacterium thermoautotrophicum* and *Methanococcus thermolithotrophicus* on ammonium chloride and dinitrogen. Arch Microbiol 148:128–131. https://doi.org/10.1007/BF00425360.

83. Pontes MH, Sevostyanova A, Groisman EA. 2015. When too much ATP is bad for protein synthesis. J Mol Biol 427:2586–2594. https://doi.org/10.1016/j.jmb.2015.06.021.

84. Torrent M, Chalancon G, de Groot NS, Wuster A, Madan Babu M. 2018. Cells alter their tRNA abundance to selectively regulate protein synthesis during stress conditions. Sci Signal 11:eaat6409. https://doi.org/10.1126/scisignal.aat6409.

85. Müller AL, Gu W, Patsalo V, Deutzmann JS, Williamson JR, Spormann AM. 2021. An alternative resource allocation strategy in the chemolithoautotrophic archaeon *Methanococcus maripaludis*. Proc Natl Acad Sci U S A 118:e2025854118. https://doi.org/10.1073/pnas.2025854118.

86. Khaova EA, Kashevarova NM, Tkachenko AG. 2022. Ribosome hibernation: molecular strategy of bacterial survival. Appl Biochem Microbiol 58:213–231. https://doi.org/10.1134/S0003683822030061.

87. Lie TJ, Dodsworth JA, Nickle DC, Leigh JA. 2007. Diverse homologues of the archaeal repressor NrpR function similarly in nitrogen regulation. FEMS Microbiol Lett 271:281–288. https://doi.org/10.1111/j.1574-6968.2007.00726.x.

88. Weidenbach K, Ehlers C, Schmitz RA. 2014. The transcriptional activator NrpA is crucial for inducing nitrogen fixation in *Methanosarcina mazei* Gö1 under nitrogen-limited conditions. FEBS J 281:3507–3522. https://doi.org/10.1111/febs.12876.

89. Prasse D, Schmitz RA. 2018. Small RNAs involved in regulation of nitrogen metabolism. Microbiol Spectr 6. https://doi.org/10.1128/microbiolspec.RWR-0018-2018.

90. Bower CE, Holm-Hansen T. 1980. A salicylate-hypochlorite method for determining ammonia in seawater. Can J Fish Aquat Sci 37:794–798. https://doi.org/10.1139/f80-106.

91. Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:113. https://doi.org/10.1186/1471-2105-5-113.

92. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Mol Biol Evol 35:1547–1549. https://doi.org/10.1093/molbev/msy096.

93. Tamura K, Stecher G, Kumar S. 2021. MEGA11: Molecular Evolutionary Genetics Analysis version 11. Mol Biol Evol 38:3022–3027. https://doi.org/10.1093/molbev/msab120.

94. Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425. https://doi.org/10.1093/oxfordjournals.molbev.a040454.

95. Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8:275–282. https://doi.org/10.1093/bioinformatics/8.3.275.

96. Kopylova E, Noé L, Touzet H. 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 28:3211–3217. https://doi.org/10.1093/bioinformatics/bts611.

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. Tamames J, Puente-Sánchez F. 2018. SqueezeMeta, a highly portable, fully automatic metagenomic analysis pipeline. Front Microbiol 9:3349. https://doi.org/10.3389/fmicb.2018.03349.

99. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550. https://doi.org/10.1186/s13059-014-0550-8.

100. Martín-Platero AM, Valdivia E, Maqueda M, Martínez-Bueno M. 2007. Fast, convenient, and economical method for isolating genomic DNA from lactic acid bacteria using a modification of the protein "salting-out" procedure. Anal Biochem 366:102–104. https://doi.org/10.1016/j.ab.2007.03.010.

101. Kolmogorov M, Bickhart DM, Behsaz B, Gurevich A, Rayko M, Shin SB, Kuhn K, Yuan J, Polevikov E, Smith TPL, Pevzner PA. 2020. metaFlye: scalable long-read metagenome assembly using repeat graphs. Nat Methods 17:1103–1110. https://doi.org/10.1038/s41592-020-00971-x.

102. Hyatt D, Chen GL, 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.

103. Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2016. GenBank. Nucleic Acids Res 44:D67–72. https://doi.org/10.1093/nar/gkv1276.

104. Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, Rattei T, Mende DR, Sunagawa S, Kuhn M, Jensen LJ, von Mering C, Bork P. 2016. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. Nucleic Acids Res 44:D286–293. https://doi.org/10.1093/nar/gkv1248.

105. Kanehisa M, Goto S. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 28:27–30. https://doi.org/10.1093/nar/28.1.27.

106. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. Nat Methods 12:59–60. https://doi.org/10.1038/nmeth.3176.

107. Eddy SR. 2009. A new generation of homology search tools based on probabilistic inference. Genome Inform 23:205–211.

108. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. 2016. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res 44:D279–285. https://doi.org/10.1093/nar/gkv1344.

109. Hallgren J, Tsirigos KD, Pedersen MD, Almagro Armenteros JJ, Marcatili P, Nielsen H, Krogh A, Winther O. 2022. DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. bioRxiv. https://10.1101/2022.04.08.487609.

110. Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791. https://doi.org/10.1111/j.1558-5646.1985.tb00420.x.