

ARTICLE OPEN

Check for updates

Phylogeny and biogeography of the algal DMS-releasing enzyme in the global ocean

Adva Shemi^{1,4}, Shifra Ben-Dor ^{2,4}, Ron Rotkopf ², Orly Dym³ and Assaf Vardi ¹

© The Author(s) 2023

Phytoplankton produce the volatile dimethyl sulfide (DMS), an important infochemical mediating microbial interactions, which is also emitted to the atmosphere and affecting the global climate. Albeit the enzymatic source for DMS in eukaryotes was elucidated, namely a DMSP lyase (DL) called Alma1, we still lack basic knowledge regarding its taxonomic distribution. We defined unique sequence motifs which enable the identification of DL homologs (DLHs) in model systems and environmental populations. We used these motifs to predict DLHs in diverse algae by analyzing hundreds of genomic and transcriptomic sequences from model systems under stress conditions and from environmental samples. Our findings show that the DL enzyme is more taxonomically widespread than previously thought, as it is encoded by known algal taxa as haptophytes and dinoflagellates, but also by chlorophytes, pelagophytes and diatoms, which were conventionally considered to lack the DL enzyme. By exploring the *Tara* Oceans database, we showed that DLHs are widespread across the oceans and are predominantly expressed by dinoflagellates. Certain dinoflagellate DLHs were differentially expressed between the euphotic and mesopelagic zones, suggesting a functional specialization and an involvement in the metabolic plasticity of mixotrophic dinoflagellates. In specific regions as the Southern Ocean, DLH expression by haptophytes and diatoms was correlated with environmental drivers such as nutrient availability. The expanded repertoire of putative DL enzymes from diverse microbial origins and geographic niches suggests new potential players in the marine sulfur cycle and provides a foundation to study the cellular function of the DL enzyme in marine microbes.

ISME Communications; https://doi.org/10.1038/s43705-023-00280-2

INTRODUCTION

Phytoplankton and marine bacteria produce the climatic gas dimethyl sulfide (DMS), a volatile organosulfur commonly known as the smell of the ocean [1]. The DMS flux from the ocean to the atmosphere is estimated as \sim 27 Tg S per year [2], making the oceans the main source of atmospheric DMS. Once emitted, DMS is being oxidized to sulfate aerosols [3], thus promoting cloud formation and increasing the albedo of Earth [4]. DMS also serves as a ubiquitous infochemical (a chemical cue that conveys information), mediating diverse trophic interactions including predator–prey interaction [5, 6], pathogenicity [7], and habitat selection [8].

The cellular precursor of DMS is dimethylsulfoniopropionate (DMSP), an osmolyte and a cryoprotectant which is found in high cellular concentration in dinoflagellates and haptophytes, and in lower cellular concentration in diatoms [9, 10]. DMSP is catabolized to DMS by distinct enzymes in bacteria and eukaryotes. The bacterial *ddd* (DMSP-dependent DMS) genes such as *dddL*, *dddD*, and *dddP*, are diverse and different from their eukaryotic counterparts [11]. In eukaryotes, a DMSP lyase (DL) enzyme called Alma1, isolated from *Emiliania huxleyi* (haptophyta), cleaves DMSP to produce DMS and acrylate. Alma1 is a redox-sensitive enzyme and a member of the Asp/Glu/Hydantoin racemase superfamily [1]. This superfamily consists of mainly bacterial or archaeal enzymes, which catalyze the conversion of enantiomers of amino acids and play important roles in the

biosynthesis of peptidoglycans and other bacterial polymers. Asp/ Glu/Hydantoin racemase genes are commonly transferred to eukaryotes by horizontal gene transfer (HGT) [12], but their biological role in the receiving organisms is currently unknown [13].

The current study provides a new perspective on the ecological significance of the eukaryotic DL enzyme and its metabolic product, DMS, in marine ecosystems. As DMS is a climatic volatile, and in light of recent elevated emissions which were linked to the current global warming [14, 15], it is crucial to deeply understand the molecular basis for DMS production. Although the ecological role of the DMS-generating enzyme in the marine sulfur cycle is widely recognized, we still lack basic knowledge regarding its cellular function [1]. Likewise, the taxonomic distribution of the DL enzyme is understudied, which hinders our ability to identify the main players in the marine sulfur cycle and to predict how environmental conditions will modulate DMS production. Following the recent discovery of the algal DMSP lyase enzyme [1], we are now able to address these key questions. We present an extensive mapping of the phylogeny and biogeographic prevalence of algal DL homologs (DLHs) by analyzing hundreds of genomic and transcriptomic sequences from model systems under diverse stress conditions in the lab and from global environmental samples. Our findings show that the DL enzyme is more taxonomically widespread than previously thought, as it is

¹Department of Plant and Environmental Sciences, Weizmann Institute of Science, Rehovot 7610001, Israel. ²Department of Life Sciences Core Facilities, Weizmann Institute of Science, Rehovot 7610001, Israel. ³Structural Proteomics Unit, Faculty of Biochemistry, Weizmann Institute of Science, Rehovot 7610001, Israel. ⁴These authors contributed equally: Adva Shemi, Shifra Ben-Dor. ^{Elemail:} assaf.vardi@weizmann.ac.il



Fig. 1 Conserved domain and sequence motifs in the DMSP lyase (DL) protein. A The DL enzyme Alma1 cleaves DMSP to produce DMS and acrylate. It has an Asp/Glu/Hydantoin conserved domain, which contains ten unique amino acid motifs. The motifs are numbered according to their order in the sequence and are colored according to their E-value score. The E-value represents how the sequence is significantly shared between 151 DLHs and also discriminant from 917 bacterial racemase domains. **B** The sequence logos of the ten identified conserved motifs, numbered as in (**A**). *According to the *E. huxleyi* Alma1 sequence [1].

encoded by known algal taxa as haptophytes and dinoflagellates, but also by chlorophytes, pelagophyte and diatoms, which were traditionally considered to lack DL. By exploring the *Tara* Oceans gene dataset [16], we demonstrate that DLH mRNA expression is occurring in diverse oceanic provinces, predominantly by dinoflagellates. Haptophyte and diatom DLH transcription was enriched in specific oceanic "hotspots", which may be affected by nutrient concentrations. The current updated taxonomy and environmental distribution of the DL enzyme opens a new avenue to investigate its yet unknown biological role(s) in phytoplankton.

RESULTS

2

Domain architecture of DLHs and their identification in diverse algal species

In order to investigate the phylogeny of the DL enzyme, we utilized the amino acid sequences from two orthologues which have high DL enzymatic activity (~600–800 μ M DMS h⁻¹): *E. huxleyi* (herein Alma1) and *Symbiodinium* Clade A1 (herein Sym-Alma) [1]. Alma1 and Sym-Alma have similar modeled 3D structures, except Sym-Alma has three cysteine residues located at the active site of the enzyme, as compared with only two cysteines in Alma1 [1] (Fig. S1). We used the Alma1 and Sym-Alma amino acid sequences, which are quite different (26% identity), as query in several databases and published genomes (see Methods), and identified a total of 153 eukaryotic DLHs (Datasets S1, S2). DLHs were defined according to their Asp/Glu/Hydantoin racemase domain (cl00518 in the CDD database, herein "racemase").

domain" [17]), and two canonical cysteine residues (C108 and C265 in Alma1), which are critical for DL activity [1] (Fig. S1). These cysteines are located in the active site and catalyze epimerisation by proton abstraction and addition. Accordingly, mutations in these residues compromised activity [1]. Racemase domain duplication occurs in 33% of identified DLHs, and a DEP domain (Disheveled, Eql-10 and Pleckstrin domain, cl02442) is conserved in several homologs (Fig. S2). Protein alignment of 151 racemase domains from eukaryotic DLHs and 34 bacterial sequences belonging to the closest racemase subfamily showed a clear difference between eukaryotic and prokaryotic homologs (which do not possess DL activity, PRK07475 in the CDD), which proposes different functions (Fig. S3). We also defined ten conserved and discriminatory sequence motifs by comparative analysis [18] between the racemase domain of all DLHs with the racemase superfamily available in the CDD database (which are bacterial sequences, see Fig. 1 and Table S1). Following confirmation of the racemase domain, the canonical cysteines, and the unique motifs, we defined 110 DLHs in 41 photosynthetic and one heterotrophic protist (Crypthecodinium cohnii, see dataset S1, S2, and Fig. S4). Lastly, corals are the only animals suggested so far to encode DLHs, which expanded to gene families in some species. The coral DLHs that were defined here were not studied in depth, as they were recently described elsewhere [19, 20].

Phylogeny of DMSP lyase homologs (DLHs)

Marine phytoplankton evolved through multiple endosymbiosis events. Primary endosymbiosis gave rise to glaucophytes,

phytes, and diatoms [21]. Photoautotrophs originated from primary endosymbiosis (including land plants, which evolved from a chlorophyte ancestor) are generally considered to lack DL enzymes, except for the green macroalga Ulva mutabilis (sea lettuce), which has two predicted DLHs [22]. Other Ulva species were shown to produce DMS, but their putative DLHs remain to be identified, since they are still missing sequenced genomes [23]. Here, six DLHs were predicted in Ulva prolifera, and new DLHs were defined in green microalgae including: Cymbomonas tetramitiformis, Tetraselmis striata, Tetraselmis suecica, and Platymonas subcordiformis, with the latter previously shown to synthesize DMSP [24] (Dataset S1). Targeted search for DLHs in additional genomes derived from the green lineage (including plants), as well as in all available glaucophytes and rhodophytes genomes, did not yield any significant putative homologs, according to the above-mentioned criteria (Dataset S3).

Photoautotrophs originated from secondary endosymbiosis include major DMS/P producers such as haptophytes and dinoflagellates [1]. We previously identified DLHs in bloomforming haptophytes as Prymnesium polylepis (formerly called Chrysochromulina polylepis), Prymnesium parvum, Isochrysis sp., Phaeocystis antarctica and E. huxleyi [1] (Table S2). Here, new homologs were identified for Isochrysis sp. and P. antarctica (A total of four and eight homologs, respectively). In addition, two DLHs were identified in a new haptophyte genome, Ochrosphaera neopolitana (Dataset S1). In contrast to the apparent gene families predicted in those haptophytes, no DLHs were detected in other haptophyte genomes, suggesting that DL conservation is not a general haptophyte trait (as Isochrysis galbana, Dataset S3). Therefore, reports for DMS-production in some haptophytes lacking DL may be derived from an alternative enzymatic source, which is yet to be discovered, or their associated bacteria [25].

In dinoflagellates, more than nine species were previously described to encode for DLHs, including *Symbiodinium* and the toxic, bloom-forming species *Karenia brevis* and *Alexandrium temarense* [1, 20]. We expanded DL gene families in most species by identifying new members, with up to 11 DLHs in *Glenodinium foliaceum* (Dataset S1). DLHs were also identified in nine new dinoflagellates: *Cladocopium goreaui, Amoebophrya ceratii, Heterocapsa triquetra*, and six *Symbiodinium* species. The various *Symbiodinium* species also have multiple DL family members. Note that *S. microadriaticum* strain CCMP2467 has sequences identical to the ones previously identified (*Symbiodinium* Clade A1) [1]. Like haptophytes, the DL gene is not common to all dinoflagellates but encoded by specific dinoflagellate genomes (Dataset S1, S3).

Unlike haptophytes and dinoflagellates, diatoms are considered to have a low cellular DMSP pool [26], and were not reported so far to have DL enzymatic activity. We identified putative DLHs in 8 ecologically important diatoms, including four Skeletonema species, a cosmopolitan bloom-forming alga [27] that can produce high DMSP levels [28]; Seminavis robusta [29]; Attheya; and the bloom-forming Pseudo-nitzschia multiseries and Pseudo-nitzschia fradulenta (Dataset S1). In addition, three putative DLHs were defined in a related group named pelagophyceae (Pelagophyceae sp. CCMP2097). A phylogenetic tree of the DL protein was constructed by aligning and comparing the racemase domain of 173 DL sequences from 34 eukaryotic and 20 bacterial species (Fig. 2, Dataset S2). The bacterial Alma1-like homologs were clearly different form their eukaryotic counterparts and appeared in the tree as an out-group. The general distribution of the species in the tree differs from their taxonomy, which may be the result of HGT from eukaryotic or bacterial sources [12, 13] (Figs. 2 and S5).

In order to validate the expression of predicted DLs under environmental conditions and to gain insights into their possible physiological function, we searched available algal transcriptomes for DLH expression in datasets generated from cells exposed to diverse environmental stress conditions [30]. Transcription of most DLHs was confirmed, including differential expression in some species in response to environmental stressors such as high light or nutrient limitation (Tables S3, S4). For example, *Skeletonema dohrnii* and *S. robusta* DLH expression was reduced by ~8–9 fold upon nitrogen depletion. In addition, in *Isochrysis* sp., expression of one out of four identified DLHs (DLH C) was increased by ~300-fold in response to high light (Table S4). Interestingly, DLH expression was species-specific under diverse environmental conditions. Future biochemical studies are required to determine DL enzymatic activity of these predicted homologs and to validate their physiological role under these stress conditions.

Global biogeographic distribution of DMSP lyase homologs in the ocean

To date, atmospheric and seawater DMS concentrations are measured on a global scale [2], yet investigation of the algal origin of DMS formation is restricted to specific ecological niches or bloom events [31, 32]. Here, the global biogeography of DLHs was examined using the Tara Oceans dataset, a collection of panoceanic plankton metagenomes and metatranscriptomes [16]. When using the DL sequence from Symbiodinium (Sym-Alma) as query (which is more similar to other DLHs, as compared to Alma1, see Dataset S1) high abundance of eukaryotic DLHs was detected in both metagenome and metatranscriptome databases (Marine Atlas of Tara Ocean Unigenes, MATOU [33]; Figs. 3 and S6). Transcripts homologous to DL were detected in all 581 samples collected from surface, the deep chlorophyl maxima and, unexpectedly, in some stations also in mesophotic depths. DLH transcripts were widely distributed across the oceans, with highest abundances in the Norwegian Sea (station Tara_163), Barents Sea (Tara_168), South Atlantic Ocean near Argentina (Tara_82) and Antarctica (Tara_85), and the North Pacific Ocean near Hawaii (Tara_131 and 132) (Fig. 3A). These locations include nutrient-enriched areas with seasonal blooms of DMS-producers such as Phaeocystis [34]. In addition, some of those sampling locations were also reported with high monthly average DMS concentration in surface water, as compared with the global annual mean concentration (~2.26 nM [2], see Table S5). The taxonomy of DLHs in those stations with highest transcript abundance was characterized as follows: DLHs were mainly from dinoflagellates such as Gonyaulacales (for example: Azadinium and Alexandrium), suessiales (Symbiodinium), peridiniales, gymnodiniales and others. Stations Tara 82 and 163 were enriched with haptophyte DLHs as Isochrysidales (E. huxleyi and Isochrysis), Phaeocystales (P. antarctica), and Prymnesiales (Prymnesium and Haptolina). Furthermore, stations Tara_173, 175, 188, and 189 were enriched with DLHs from diatoms, mainly Thalassionematales (Thalassiothrix antarctica), Thalassiosirales (Thalassiosira antarctica) and Chaetocerotales (as Chaetoceros, which was previously shown to produce DMSP [35]) (Figs. 3B and S7). Heterotrophs such as radiolarians, ciliates and metazoans, which often harbor photosymbionts, were also associated with DLH expression (Fig. 3B, "other"). Interestingly, those holobionts are considered as a significant source for DMS/P [19, 36].

To further investigate the environmental drivers of DL mRNA expression, which hint at physiological roles for DL, we analyzed the correlation between abiotic factors measured in each station (temperature and nutrients) and the relative abundance of DLH sequences in all metatranscriptomes versus metagenomes (Fig. 3C). Dinoflagellate DLH mRNAs were generally highly expressed across the oceans, regardless of the environmental parameters tested. In the South Pacific and Southern Oceans, haptophyte DLH mRNA expression (mainly *Phaeocystis*) was positively correlated with nutrient availability (nitrogen, phosphate, or iron). This contrasts with diatom-derived DLH expression in the Southern Ocean, which correlated with low nutrients, suggesting that the DL cellular role under stress conditions may be taxon-specific.



Fig. 2 Phylogenetic distribution of the DMSP lyase enzyme. A phylogenetic tree of identified DLHs, colored according to their taxonomy. Gray background designates experimental evidence for mRNA expression (see Tables S3, S4). Red arrows indicate biochemical confirmation of DL activity in the marked homologs. The lsochrysis sp. orthologs are \geq 99% identical to alma1 from *E. huxleyi* and are therefore confirmed by identity [1].

Dinoflagellates are unique phytoplankton since many species inhabit the sunlit as well as dark deep ocean, owing to their versatile metabolism which alternates between photoautotrophy and heterotrophy (also known as mixotrophy [37]). As dinoflagellate DLH expression was not affected by the environmental parameters measured during the *Tara* campaign, we asked whether they differentially express DL in surface samples versus deep water samples. Indeed, a remarkable abundance of dinoflagellate DLH transcripts was detected in the deep dark ocean (200–740 m) in 11 *Tara* stations (Fig. S8). We further explored this phenomenon in a metatranscriptomic study presenting a detailed vertical profile of dinoflagellate populations from the surface to the depth (800 m) of the central Pacific Ocean (the METZYME transect [38]). In addition to the light gradient, deep water was rich in nitrogen, which was gradually depleted toward the surface. Dinoflagellate DLH transcripts were prevalent throughout the water column. Among 23 distinct dinoflagellate DLHs identified, 16 homologs were significantly expressed in the euphotic, nitrogen-limited zone ($\leq 200 \text{ m depth}$). Interestingly, four homologs were significantly upregulated in the dark mesopelagic, nitrogen-replete zone ($\geq 200 \text{ m}$; two from *H. triquetra*, one from *Gambierdiscus polynesiensis* and one from *Durinskia baltica*, similarity >83%, Fig. S9). Taken together, the differential expression of dinoflagellate DLHs between the euphotic and mesopelagic zones suggests functional specialization of DLHs, which can act as part of the metabolic plasticity in

4



5

Fig. 3 Biogeographic distribution of DMSP lyase homologs (DLHs) in the global ocean. A The relative abundance of DLH transcripts in metatranscriptomes from 86 locations worldwide, in samples collected from surface water (5 m). The circles are proportional to transcript relative abundance (normalized as percent of mapped reads) and colored according to the different size fractions collected for each sample. The numbers designate *Tara* stations which were further analyzed in (**B**). **B** Taxonomic distribution of DLHs in the *top 20 stations with the highest DLHs expression, as described in (**A**). **C** Correlation between environmental conditions and total DLH abundance in 316 DNA samples (left panel) and 401 RNA samples (right panel). Pairwise comparisons are shown with a color gradient denoting Spearman's correlation coefficient. Comparisons were done per taxa and per each oceanic region. For metagenomes, n = 60 for Indian; 83 for North Atlantic; 65 for South Atlantic; 64 for North Pacific; 32 for the South Pacific; and 12 for the Southern Ocean. For metatranscriptomes, n = 66 for Indian and North Pacific; 95 for North Atlantic; 71 for South Atlantic; 90 for South Pacific; and 14 for the Southern Ocean. The DLHs abundance and taxonomic distribution were based on the 0.8–2000 µm size fraction from all depths available in the *Tara* Oceans dataset. Dino. Dinophyceae, Hapt. Haptophyceae, Diat. Diatoms, Euka. Other eukaryotes. *Estimated values based on oceanographic models [16, 65].

mixotrophic dinoflagellates [38]. In summary, DLH transcript abundance in natural populations was correlated to nutrient and/or light availability in a taxon-specific manner, thus suggesting a function for the DL enzyme as part of nutrient and carbon metabolism and the recycling machinery in phytoplankton.

DISCUSSION

Oceanic DMS is an important component of the marine sulfur cycle, with ecological implications ranging from chemical signaling across trophic levels to the global climate. Here, we investigated the phylogeny and geographic distribution of the eukaryotic DMS-releasing enzyme, the DL Alma1 [1]. DLHs were



B Ocean-atmosphere interface Sulfate aerosols → Cloud formation



detected in the genomes of dinoflagellates, haptophytes, diatoms, chlorophytes, pelagophytes, and scleractinians. The expanded repertoire of putative DL enzymes from diverse microbial origins suggests new potential players in the marine sulfur cycle beyond the known representatives from dinoflagellates and haptophytes. We revealed that DLHs are encoded in the genomes of distantly related taxa (Dataset S1, Fig. 2), and could not detect DLH in red algal genomes. Interestingly, no known DLH was identified in Polysiphonia, from which DMS was historically detected for the first time [39]. Thus, DMS production by this red seaweed may be attributed to alternative enzymatic sources or to strain variability (according to available genomes). Therefore, DL enzymes in groups possessing secondary plastids as dinoflagellates, haptophytes, diatoms, and pelagophytes, may have originated from a green ancestor [40], or from the heterotrophic host or red alga (prior to gene loss). In addition, the distribution of DLHs in the

SPRINGER NATURE

Fig. 4 The evolution, cellular and ecological roles for the DMSP lyase enzyme in phytoplankton. A Possible acquisition pathways for the DL enzyme in haptophytes, dinoflagellates, diatoms and pelagophytes. Primary endosymbiosis gave rise to chlorophytes, including species which possess DLHs, and to rhodophytes and glaucophytes, which their DLH are either missing or remained to be identified. Secondary endosymbiosis gave rise to haptophytes, dinoflagellates, diatoms and pelagophytes, which all have representative species encoding for DLHs. The DL evolutionary origin in those phytoplankton groups may be attributed to the heterotrophic host that engulfed a rhodophyte or chlorophyte, to a rhodophyte (considering unidentified DL, or that DL gene loss in red genomes occurred after secondary endosymbiosis), or to chlorophytes [40] or bacteria via horizontal gene transfer (HGT) [12]. Corals DLHs were probably horizontally transferred from their dinoflagellate symbionts [19, 20]. B Proposed cellular and ecological roles for the DL enzyme under stress conditions. DL enzymatic activity or transcription was linked (up- or down-regulated, depending on tested species) to nitrogen levels (this study, and ref. [29, 46]); high light and ultraviolet (UV) radiation (this study, and ref. [42, 44, 45]; cold stress [22]; carbon dioxide [42], and oxidative stress [1]. During microbial interactions, DMS acts as an activation cue for parasites [7] and plays a pro-grazing role for protist grazers [6] and copepods [51]. DMS is also accumulated in the water during grazing, serving as chemoattractant for diverse top predators [66]. In addition to the mentioned cellular and ecological roles, DMS plays a climatic role; the ocean-atmosphere DMS flux is estimated as 10 million tons S per year [2], and atmospheric DMS is rapidly oxidized to form sulfate aerosols, enhancing cloud formation and increases the albedo of earth.

phylogenetic tree suggests that HGTs from prokaryotic and eukaryotic genomes may have played a role during the evolution of the DL enzyme [12, 13] (Figs. 2 and 4A).

The prevalence of DLHs in diverse, but highly selective, algal genomes, proposes important cellular role(s) for this enzyme. Unlike DMSP, which is a known osmolyte [41], cryoprotectant, antioxidant [42], and energy dissipation metabolite [43], very little is known regarding the biological function of the DL enzyme. Does the DL activity act in controlling the cellular DMSP levels? Does DMS play a cellular role as well? Previous studies had shown that temperature, high light [44, 45], carbon dioxide [42], and nitrogen [46] modulate DMS production in different phytoplankton, but the DMS role is still not fully understood (Fig. 4B). Such environmental stress conditions can induce cellular reactive oxygen species (ROS) levels, which can be a mechanism for DL activation or inhibition, as it is a redox-sensitive enzyme [1]. Furthermore, DMS can scavenge ROS by being oxidized to dimethyl sulfoxide, thus acting as a potent antioxidant and promoting cellular resilience to stress [42, 47].

DLH expression by most identified taxa (Fig. 2) was detected across the oceans, with dinoflagellates emerging as the dominant group expressing DLHs, regardless of environmental conditions. This suggests that the DL enzyme may be part of the previously described dinoflagellate "core" component of genes with constitutive expression or little differential expression [48]. Expression of dinoflagellate DLHs even detected in mesopelagic depths, implying a possible novel metabolic role for DLH in mixotrophy [38]. Another option is that DMS generation by those mixotrophs is a strategy to attract algal or bacterial prey [49]. Haptophytes, diatoms and others expressed DLHs in correlation with some nutrient availability in specific oceanic provinces, suggesting a link between DMS/P production and nutrient metabolism which is still requires further investigation [46] (Fig. 4A). For instance, high DLH expression by diatoms was correlated with low nitrogen and iron levels in the Southern Ocean; this can potentially be triggered by increase in cellular DMSP, as was previously shown for nitrogen-depleted Skeletonema [28] and nitrogen- or iron-depleted Thalassiosira [26, 50] (Fig. 3D). This potential vital role for the DL enzyme in diverse phytoplankton under ecologically-relevant conditions is especially ivity constrongly Identi

interesting, as it may come with a cost, since DL activity can strongly enhance predation by zooplankton [6, 51] (Fig. 4B). Consequently, there should be high negative selection to retain DMS-generating enzymes in algal genomes. It is plausible that a tradeoff between abiotic stress conditions (such as nutrient limitation) and grazing pressure promotes the selective retention of the DL enzyme in phytoplankton, including ecologically successful species such as *E. huxleyi* and *Symbiodinium* (Fig. 4B).

Elucidating the DL function in its physiological context and during microbial interactions poses great methodological challenges. Future research should confirm DL enzymatic activity in cultured species, which will extend the repertoire of model organisms available to study the DL physiological role in phytoplankton. To date, specific in vitro DL activity was demonstrated by expression and purification of recombinant DL enzymes only for E. huxleyi and Symbiodinium Clade A1 [1]. In Phaeocystis, Ulva and others, DMS was detected in cell extracts and not from a specific gene product [1, 52, 53]. Intriguingly, the reef-building coral Acropora encodes a set of genes similar to Alma1 and Sym-Alma (~43% and 60% similarity in the amino acid sequence, respectively. Dataset S1 and [20]). DLH genes were predicted to be horizontally transferred to the coral genome from its dinoflagellate symbiont [19], and their enzymatic activity remains to be fully validated [54]. However, the absence of genetic tools to interrogate DL activity and gene function in DMS-producing species is a major bottleneck hampering our understanding of the DL eco-physiological role. The current study provides the basis for developing model systems and experimental setups needed to fulfill this goal. We identified DLHs in several model species that are amenable to genetic transformation as Ps. multiseries, T. striata, C. cohnii, Amphidinium carterae [55] and Ulva [56]. Those can be mutated in the DL conserved sequence motifs and active site residues (Fig. 1), followed by activity assay to validate the predicted DL enzymes. In addition, heterologous overexpression of the predicted sequences from various species (shown in Fig. 2) in Escherichia coli, including point mutations in key amino acids and testing of DL activity, will confirm the function of putative DL homologs [1].

Furthermore, we demonstrated wide distribution of DLH transcripts across the oceans and found possible links between haptophyte and diatom DLHs expression and environmental conditions in specific oceanic regions (Fig. 3). This 'snapshot' of gene expression patterns available from the *Tara Ocean* dataset can provide the base for hypothesis-driven, time course experiments under controlled laboratory conditions, to test for the DL role in response to abiotic stress (For example, during acclimation to nitrogen limitation). In addition, applying a selective DL inhibitor [54] during growth experiments under stress conditions and in the course of biological interactions [6, 7] could shed light on DL cellular function.

With the apparent impact of the global climate change on DMS emissions [14, 15], there is a critical need to understand the biological origins of DMS in the ocean. Our findings have expanded the taxonomic and biogeographic distribution of the DL enzymes, as well as its potential cellular roles. We provide new genomic resources and environmental insights that will help to unravel the ecological importance of this unique enzyme in oceanic ecosystems.

MATERIALS AND METHODS

Predicted protein structures of DL orthologs

In order to predict the structure of the DL enzyme, a BLAST [57] search was conducted against the Protein Data Bank (PDB) using the DL sequences from *Symbiodinium* (Sym-Alma) and *E. huxleyi* (Alma1) as queries. Since no significant similarity to any known protein structure was found, theoretical structural models were created using the AlphaFold server [58, 59]. Ribbon representation (Fig. S1) was created using PyMOL (The PyMOL Molecular Graphics System, v. 1.2r3pre, Schrödinger, LLC).

Identification of DLHs

In order to identify DLHs, database searching was performed at various websites using BLAST [57], either BLASTP or TBLASTN, with three sequences as initial input: E. huxleyi Alma 1 (AKO62592), Symbiodinium (PODN22) and Acropora millepora (XP_029211597), and using an E-value threshold of E-05. Data sources searched for DLHs include the Marine Microbial Eukaryotic Transcriptome Sequencing Project (MMETSP) [30]; the Joint Genome Institute (JGI) Phycocosm; Supplemental data from Nelson et al. [60]; and various genomes and TSA in NCBI. For red algal genomes, the "Red Algal Resources to Promote Integrative Research in Algal Genomics" Blast site [61] was used. For Ac. millepora, the draft genome (v. 2.01) published by the Przeworski lab [62] was used. When a sequences from species close to the target had already been defined, those sequences were also utilized as gueries in the searches. For downloaded genomes and datasets, local BLAST [57] was performed (v. 2.11.0). Genes were constructed manually as described in Feldmesser et al. [63]. Sequences found in the initial search were then studied further, and considered as DLHs only if they contained both canonical cysteines at the active site (C108 and C265 from E. huxleyi Alma1) [1] and the surrounding sequences (DCGF or NCGF and ECTE or ECTQ, respectively). Domain analysis was performed on the sequences using the Batch CD-Search (CDD v3.19) at NCBI [17]. Only sequences with a hit to Asp/Glu/Hydantoin racemase superfamily domain (cl00518) were retained.

DL sequence motifs and logos

To define unique sequence motifs for the DL enzyme, differential motifs in the amino acid sequence were analyzed with Streme (v. 5.4.1), a part of the MEME suite, using 151 DLH sequences without bacterial or putative bacterial sequences from the alignment. A total of 917 bacterial sequences from the racemase domain family in the CDD were used as control. Default parameters were applied except for logo length, where 4–15 positions were allowed. Logos of the alignment were constructed with WebLogo (v. 3.0).

DLHs sequence alignments and phylogenetic tree

In order to align the identified DLHs for phylogenetic analysis, sequences were cut to the domain (according to the CDD search), and alignments were performed using Muscle (v. 3.8.31) and ClustalW (v. 2.1) [64]. The final alignment (Muscle) for further analysis was cleaned of extensions by removal of all columns with less than 10 sequences. The alignment was used as input for similarity/identity calculation using MacVector (v. 18.0). Phylogenetic tree was constructed using Neighbor Joining in ClustalW with 1000 bootstraps (v. 2.1) and Maximum Likelihood with the default parameters in PhyML (v. 3.0). The PhyML trees were visualized with iTol (v. 6).

Transcription of DLHs in phytoplankton cultures

To assess if predicted DLHs are expressed, mRNA expression level in cultured cells under abiotic stress conditions was obtained from the MMETSP database [30] (Tables S3, S4). Additional transcriptomic studies are referenced in Table S4. Read counts were normalized with Deseq2. Expression data were also taken from metatranscriptomic studies, as described below.

Biogeography of DLHs in the Tara Ocean dataset

In order to map the bio-geographic distribution of DLHs in the marine environment, the Ocean Gene Atlas platform was used [16, 65]. DLHs search was performed with TBLASTN, utilizing *Sym*-Alma (from *Symbiodinium* Clade A1 [1]) as query. The databases queried were the Tara Oceans Eukaryotic Genomes (SMAGs) and the Marine Atlas of Tara Oceans Unigenes (MATOU_v1_metaG, eukaryotes) for metagenomes, and the MATOU_v2_metaT for metatranscriptomes. The expected threshold was 1E-10. The abundance was calculated as percent of mapped reads.

Dinoflagellate DLHs expression from the METZYME transect

To compare dinoflagellates DLHs expression between euphotic and mesophotic depths, we searched for DLHs in the metatranscriptomic dataset published by Cohen et al. [38] using the Blast criteria listed above (TBLASTN, conserved cysteines and surrounding sequences). The putative DLH expression was then extracted from the dataset.

Illustrations were created with BioRender.com.

Graphs were created with GraphPad Prism.

8

DATA AVAILABILITY

The data used in this study is available in datasets S1-S3.

REFERENCES

- Alcolombri U, Ben-Dor S, Feldmesser E, Levin Y, Tawfik DS, Vardi A. Identification of the algal dimethyl sulfide-releasing enzyme: a missing link in the marine sulfur cycle. Science. 2015;348:1466–9.
- Hulswar S, Simó R, Galí M, Bell TG, Lana A, Inamdar S, et al. Third revision of the global surface seawater dimethyl sulfide climatology (DMS-Rev3). Earth Syst Sci Data. 2022;14:2963–87.
- Simó R. Production of atmospheric sulfur by oceanic plankton: biogeochemical, ecological and evolutionary links. Trends Ecol Evol. 2001;16:287–94.
- Charlson RJ, Lovelock JE, Andreae MO, Warren SG. Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate. Nature. 1987;326:655–61.
- Owen K, Saeki K, Warren J, Bocconcelli A, Wiley D, Ohira S-I, et al. Natural dimethyl sulfide gradients would lead marine predators to higher prey biomass. Commun Biol. 2021;4:149.
- Shemi A, Alcolombri U, Schatz D, Farstey V, Vincent F, Rotkopf R, et al. Dimethyl sulfide mediates microbial predator–prey interactions between zooplankton and algae in the ocean. Nat Microbiol. 2021;6:1357–66.
- Garcés E, Alacid E, Reñé A, Petrou K, Simó R. Host-released dimethylsulphide activates the dinoflagellate parasitoid *Parvilucifera sinerae*. ISME J. 2013; 7:1065–8.
- Foretich MA, Paris CB, Grosell M, Stieglitz JD, Benetti DD. Dimethyl sulfide is a chemical attractant for reef fish larvae. Sci Rep. 2017;7:2498.
- Yoch DC. Dimethylsulfoniopropionate: its sources, role in the marine food web, and biological degradation to dimethylsulfide. Appl Environ Microbiol. 2002;68:5804–15.
- Caruana AMN, Malin G. The variability in DMSP content and DMSP lyase activity in marine dinoflagellates. Prog Oceanogr. 2014;120:410–24.
- Bullock HA, Luo H, Whitman WB. Evolution of dimethylsulfoniopropionate metabolism in marine phytoplankton and bacteria. Front Microbiol. 2017;8:637.
- Dougan KE, Deng Z-L, Wöhlbrand L, Reuse C, Bunk B, Chen Y, et al. Multi-omics analysis reveals the molecular response to heat stress in a "red tide" dinoflagellate. *bioRxiv*. 2022:2022.07.25.501386.
- Naranjo-Ortíz MA, Brock M, Brunke S, Hube B, Marcet-Houben M, Gabaldón T. Widespread inter- and intra-domain horizontal gene transfer of d-amino acid metabolism enzymes in eukaryotes. *Front Microbiol.* 2016;7:2001.
- Galí M, Devred E, Babin M, Levasseur M. Decadal increase in Arctic dimethylsulfide emission. Proc Natl Acad Sci USA. 2019;116:19311–17.
- Park K, Kim I, Choi J-O, Lee Y, Jung J, Ha S-Y, et al. Unexpectedly high dimethyl sulfide concentration in high-latitude Arctic sea ice melt ponds. Environ Sci Process Impacts. 2019;21:1642–9.
- Vernette C, Lecubin J, Sánchez P, Coordinators TO, Sunagawa S, Delmont TO, et al. The Ocean Gene Atlas v2.0: online exploration of the biogeography and phylogeny of plankton genes. Nucleic Acids Res. 2022;50:W516–W26.
- Marchler-Bauer A, Bryant SH. CD-Search: protein domain annotations on the fly. Nucleic Acids Res. 2004;32:W327–31.
- Bailey TL. STREME: accurate and versatile sequence motif discovery. Bioinformatics. 2021;37:2834–40.
- Shinzato C, Khalturin K, Inoue J, Zayasu Y, Kanda M, Kawamitsu M, et al. Eighteen coral genomes reveal the evolutionary origin of *Acropora* strategies to accommodate environmental changes. Mol Biol Evol. 2020;38:16–30.
- Chiu Y-L, Shinzato C. Evolutionary history of DMSP lyase-like genes in animals and their possible involvement in evolution of the scleractinian coral genus, *Acropora*. *Front Mar Sci.* 2022;9:889866.
- Chan CX, Gross J, Yoon HS, Bhattacharya D. Plastid origin and evolution: new models provide insights into old problems. Plant Physiol. 2011;155:1552–60.
- De Clerck O, Kao SM, Bogaert KA, Blomme J, Foflonker F, Kwantes M, et al. Insights into the evolution of multicellularity from the sea lettuce genome. Curr Biol. 2018;28:2921–33.e5.
- de Souza MP, Chen YP, Yoch DC. Dimethylsulfoniopropionate lyase from the marine macroalga *Ulva curvata*: purification and characterization of the enzyme. Planta. 1996;199:433–8.
- Dickson DMJ, Kirst GO. The role of β-dimethylsulphoniopropionate, glycine betaine and homarine in the osmoacclimation of *Platymonas subcordiformis*. Planta. 1986;167:536–43.
- Gebser B, Thume K, Steinke M, Pohnert G. Phytoplankton-derived zwitterionic gonyol and dimethylsulfonioacetate interfere with microbial dimethylsulfoniopropionate sulfur cycling. MicrobiologyOpen. 2020;9:e1014.
- McParland EL, Wright A, Art K, He M, Levine NM. Evidence for contrasting roles of dimethylsulfoniopropionate production in *Emiliania huxleyi* and *Thalassiosira* oceanica. New Phytol. 2020;226:396–409.

- Johansson ON, Töpel M, Pinder MIM, Kourtchenko O, Blomberg A, Godhe A, et al. Skeletonema marinoi as a new genetic model for marine chain-forming diatoms. Sci Rep. 2019;9:5391.
- Spielmeyer A, Pohnert G. Daytime, growth phase and nitrate availability dependent variations of dimethylsulfoniopropionate in batch cultures of the diatom *Skeletonema marinoi*. J Exp Mar Biol Ecol. 2012;413:121–30.
- Osuna-Cruz CM, Bilcke G, Vancaester E, De Decker S, Bones AM, Winge P, et al. The Seminavis robusta genome provides insights into the evolutionary adaptations of benthic diatoms. Nat Commun. 2020;11:3320.
- Keeling PJ, Burki F, Wilcox HM, Allam B, Allen EE, Amaral-Zettler LA, et al. The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. PLOS Biol. 2014;12:e1001889.
- Steinke M, Malin G, Gibb SW, Burkill PH. Vertical and temporal variability of DMSP lyase activity in a coccolithophorid bloom in the northern North Sea. Deep-Sea Res PT II. 2002;49:3001–16.
- Steinke M, Malin G, Archer S, Burkill P, Liss P. DMS production in a coccolithophorid bloom: evidence for the importance of dinoflagellate DMSP lyases. Aquat Microb Ecol. 2002;26:259–70.
- Carradec Q, Pelletier E, Da Silva C, Alberti A, Seeleuthner Y, Blanc-Mathieu R, et al. A global ocean atlas of eukaryotic genes. Nat Commun. 2018;9:373.
- Schoemann V, Becquevort S, Stefels J, Rousseau V, Lancelot C. Phaeocystis blooms in the global ocean and their controlling mechanisms: a review. J Sea Res. 2005;53:43–66.
- Kasamatsu N, Hirano T, Kudoh S, Odate T, Fukuchi M. Dimethylsulfoniopropionate production by psychrophilic diatom isolates. J Phycol. 2004;40:874–8.
- Gutierrez-Rodriguez A, Pillet L, Biard T, Said-Ahmad W, Amrani A, Simó R, et al. Dimethylated sulfur compounds in symbiotic protists: a potentially significant source for marine DMS(P). Limnol Oceanogr. 2017;62:1139–54.
- Jeong HJ, Yoo YD, Kim JS, Seong KA, Kang NS, Kim TH. Growth, feeding and ecological roles of the mixotrophic and heterotrophic dinoflagellates in marine planktonic food webs. Ocean Sci. 2010;45:65–91.
- Cohen NR, McIlvin MR, Moran DM, Held NA, Saunders JK, Hawco NJ, et al. Dinoflagellates alter their carbon and nutrient metabolic strategies across environmental gradients in the central Pacific Ocean. Nat Microbiol. 2021;6:173–86.
- 39. Haas P. The liberation of methyl sulphide by seaweed. Biochem J. 1935;29:1297–9.
- Moustafa A, Beszteri B, Maier UG, Bowler C, Valentin K, Bhattacharya D. Genomic footprints of a cryptic plastid endosymbiosis in diatoms. Science. 2009;324:1724–6.
- Kirst GO. Osmotic adjustment in phytoplankton and macroalgae. In: Kiene RP, Visscher PT, Keller MD, Kirst GO, editors. Biological and environmental chemistry of DMSP and related sulfonium compounds. Boston, MA: Springer US; 1996. p. 121–9.
- 42. Sunda W, Kieber DJ, Kiene RP, Huntsman S. An antioxidant function for DMSP and DMS in marine algae. Nature. 2002;418:317–20.
- Stefels J. Physiological aspects of the production and conversion of DMSP in marine algae and higher plants. J Sea Res. 2000;43:183–97.
- Archer SD, Ragni M, Webster R, Airs RL, Geider RJ. Dimethyl sulfoniopropionate and dimethyl sulfide production in response to photoinhibition in *Emiliania huxleyi*. Limnol Oceanogr. 2010;55:1579–89.
- Darroch L, Lavoie M, Levasseur M, Laurion I, Sunda W, Michaud S, et al. Effect of short-term light- and UV-stress on DMSP, DMS, and DMSP lyase activity in *Emiliania huxleyi*. Aquat Microb Ecol. 2015;74:173–85.
- Sunda WG, Hardison R, Kiene RP, Bucciarelli E, Harada H. The effect of nitrogen limitation on cellular DMSP and DMS release in marine phytoplankton: climate feedback implications. Aquat Sci. 2007;69:341–51.
- Thume K, Gebser B, Chen L, Meyer N, Kieber DJ, Pohnert G. The metabolite dimethylsulfoxonium propionate extends the marine organosulfur cycle. Nature. 2018;563:412–5.
- 48. Moustafa A, Evans AN, Kulis DM, Hackett JD, Erdner DL, Anderson DM, et al. Transcriptome profiling of a toxic dinoflagellate reveals a gene-rich protist and a potential impact on gene expression due to bacterial presence. PLOS One. 2010;5:e9688.
- Seymour J, Simó R, Ahmed T, Stocker R. Chemoattraction to dimethylsulfoniopropionate throughout the marine microbial food web. Science. 2010;329:342–5.
- Bucciarelli E, Ridame C, Sunda WG, Dimier-Hugueney C, Cheize M, Belviso S. Increased intracellular concentrations of DMSP and DMSO in iron-limited oceanic phytoplankton *Thalassiosira oceanica* and *Trichodesmium erythraeum*. Limnol Oceanogr. 2013;58:1667–79.
- 51. Procter J, Hopkins FE, Fileman ES, Lindeque PK. Smells good enough to eat: Dimethyl sulfide (DMS) enhances copepod ingestion of microplastics. Mar Pollut Bull. 2019;138:1–6.

- Steinke M, Wolfe GV, Kirst GO. Partial characterisation of dimethylsulfoniopropionate (DMSP) lyase isozymes in 6 strains of *Emiliania huxleyi*. Mar Ecol. 1998;175:215–25.
- Stefels J, van Boekel WHM. Production of DMS from dissolved DMSP in axenic cultures of the marine phytoplankton species *Phaeocystis* sp. Mar Ecol Prog Ser. 1993;97:11–8.
- Alcolombri U, Lei L, Meltzer D, Vardi A, Tawfik DS. Assigning the algal source of dimethylsulfide using a selective lyase inhibitor. ACS Chem Biol. 2017;12:41–6.
- 55. Faktorová D, Nisbet RER, Fernández Robledo JA, Casacuberta E, Sudek L, Allen AE, et al. Genetic tool development in marine protists: emerging model organisms for experimental cell biology. Nat Methods. 2020;17:481–94.
- Oertel W, Wichard T, Weissgerber A. Transformation of *Ulva mutabilis* (Chlorophyta) by vector plasmids integrating into the genome. J Phycol. 2015;51:963–79.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10.
- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596:583–9.
- Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, et al. AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. 2021;50:D439–D44.
- Nelson DR, Hazzouri KM, Lauersen KJ, Jaiswal A, Chaiboonchoe A, Mystikou A, et al. Large-scale genome sequencing reveals the driving forces of viruses in microalgal evolution. Cell Host Microbe. 2021;29:250–66.e8.
- Rutgers University, grant NPR. Red algal resources to promote integrative research in algal genomics. http://porphyra.rutgers.edu/.
- Przeworski M. Draft genome of the staghorn coral Acropora millepora. https:// przeworskilab.com/data/.
- Feldmesser E, Rosenwasser S, Vardi A, Ben-Dor S. Improving transcriptome construction in non-model organisms: integrating manual and automated gene definition in *Emiliania huxleyi*. BMC Genom. 2014;15:148.
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007;23:2947–8.
- Villar E, Vannier T, Vernette C, Lescot M, Cuenca M, Alexandre A, et al. The Ocean Gene Atlas: exploring the biogeography of plankton genes online. Nucleic Acids Res. 2018;46:W289–W95.
- Savoca MS. Chemoattraction to dimethyl sulfide links the sulfur, iron, and carbon cycles in high-latitude oceans. Biogeochemistry. 2018;138:1–21.

ACKNOWLEDGEMENTS

We thank Dr. Gust Bilcke for providing the DLH mRNA expression data for *S. robusta* (Table S4). We thank Dr. Uria Alcolombri for his contribution to scientific discussions and his constructive feedback on the manuscript. This work was funded by the

research grant from the Israeli Science Foundation (ISF No. 1972/20) and by a research grant from the Takiff Family Foundation, both were awarded to A.V.

AUTHOR CONTRIBUTIONS

AS and AV conceptualized this study and wrote the manuscript. SBD identified DLHs, defined sequence motifs and analyzed the DL phylogeny. AS analyzed the geographic distribution, taxonomy and environmental drivers of DLH expression in the Tara Oceans dataset and created the figures. SBD, AS, and RR analyzed the dinoflagellate DLH expression from the METZYME dataset. RR conducted statistical analysis and normalization of expression data. OD created the structural predictions for DL orthologs.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43705-023-00280-2.

Correspondence and requests for materials should be addressed to Assaf Vardi.

Reprints and permission information is available at http://www.nature.com/ reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http:// creativecommons.ord/licenses/by/4.0/.

© The Author(s) 2023