Title - Selective lipid recruitment by an archaeal DPANN symbiont from its host

Keywords - DPANN archaea, symbiont, lipidomics, molecular network

Authors list

Su Ding[1,†,*], Joshua N. Hamm[1,†,*], Nicole J. Bale[1], Jaap S. Sinninghe Damsté[1,2], Anja Spang[1,3]

Affiliations

[1]NIOZ Royal Institute for Sea Research, Department of Marine Microbiology and Biogeochemistry, Texel, The Netherlands

[2]Department of Earth Sciences, Faculty of Geosciences, Utrecht University, Utrecht, The Netherlands

[3]Department of Evolutionary & Population Biology, Institute for Biodiversity and Ecosystem Dynamics (IBED), University of Amsterdam, Amsterdam, The Netherlands

[†]These authors contributed equally to this work

[*]Correspondence: su.ding@nioz.nl; joshua.hamm@nioz.nl

Abstract

The symbiont *Ca.* Nanohaloarchaeum antarcticus is obligately dependent on its host *Halorubrum lacusprofundi* for lipids and other metabolites due to its lack of certain biosynthetic genes. However, it remains unclear which specific lipids or metabolites are acquired from its host, and how the host responds to infection. Here, we explored the lipidome dynamics of the *Ca.* Nha. antarcticus – *Hrr. lacusprofundi* symbiotic relationship during co-cultivation. By using a comprehensive untargeted lipidomic methodology, our study reveals that *Ca.* Nha. antarcticus selectively recruits 110 lipid species from its host, i.e. nearly two-thirds of the total number of host lipids. Lipid profiles of co-cultures displayed shifts in abundances of bacterioruberins and menaquinones and changes in degree of bilayer-forming glycerolipid unsaturation. This likely results in increased membrane fluidity and improved resistance to membrane disruptions, consistent with compensation for higher metabolic load and mechanical stress on host membranes when in contact with *Ca.* Nha. antarcticus cells. Notably, our findings differ from previous observations of other DPANN symbiont-host systems, where no differences in lipidome composition were reported. Altogether, our work emphasizes the strength of employing untargeted lipidomics approaches to provide details into the dynamics underlying a DPANN symbiont-host system.

Main text

Members of the DPANN archaea, originally named after the initials of its first five identified groups (Diapherotrites, Parvarchaeota, Aenigmarchaeota, Nanoarchaeota, and Nanohaloarchaeota), are characterized by small cell and genome sizes[1]. Since their discovery, additional lineages such as Woesearchaeota[2] and Pacearchaeota[2], Huberarchaeota[3], Micrarchaeota[4], Altiarchaeota[5], Undinarchaeota[6], and Mamarchaeota[7] have been identified and incorporated into the DPANN superphylum. These archaea are widely distributed in diverse environments, including hypersaline lakes[8, 9], marine[6, 10] and freshwater[1, 2] bodies, sediments[11, 12], acid mine drainage sites[4], and hot springs[13]. Apart from Altiarchaeota, most DPANN archaea exhibit limited catabolic capabilities necessary to sustain a free-living lifestyle[2, 14] and the majority of them are predicted to rely on symbiotic interactions with other organisms[14, 15]. Indeed, the few cultivated DPANN have symbiotic lifestyles, with representatives from three lineages (Nanoarchaeota, Nanohaloarchaeota, and Micrarchaeota) available in co-culture with specific host archaea[9, 13, 16, 17, 18, 19, 20, 21, 22, 23, 24].

Due to incomplete biosynthetic pathways for nucleotides, amino acids, and lipids, most DPANN representatives are predicted to be dependent on metabolites from their hosts. Our current knowledge regarding the identity of those metabolites and the molecular basis for the exchange and/or uptake of these compounds is limited. Two DPANN representatives (i.e. *Nanoarchaeum equitans* and *Ca.* Micrarchaeum harzensis) are thought to acquire lipids from their hosts (*Ignicoccus hospitalis*[25] and *Ca.* Scheffleriplasma hospitalis[21], respectively). Lipid analyses of pure host and symbionts as well as symbiont-host co-cultures have revealed that there is no significant qualitative difference in the lipid composition profiles between those cultures, indicating that the process of lipid uptake from the host is non-selective. However, it remains unclear whether non-selective lipid uptake from host organisms is a common feature among the various DPANN representatives.

In this study, we conducted a comprehensive analysis of the lipidome of the DPANN symbiont-host system consisting of *Ca.* Nanohaloarchaeum antarcticus – *Halorubrum lacusprofundi*[9, 26]. *Ca.* Nha. antarcticus has so far not been obtained in a stable pure co-culture with *Hrr. lacusprofundi* and must be maintained in an enrichment culture (CLAC2B) containing multiple *Hrr. lacusprofundi* strains along with a *Natrinema* sp.. Recent work has shown that the instability of pure co-cultures may be due to *Ca.* Nha. antarcticus being a parasite that invades the host cytoplasm leading to host cell lysis[26]. *Ca.* Nha. antarcticus lacks identifiable genes encoding proteins involved in lipid biosynthesis and metabolism, and is thus hypothesized to rely on lipids of its host for survival[9]. Through the investigation of the lipidome of *Ca.* Nha. antarcticus and its host, we aimed to 1) determine whether the lipids of *Ca.* Nha. antarcticus closely resemble those of the host or exhibit differences; 2) assess for potential changes in the host's membrane lipid composition upon infection by *Ca.* Nha. antarcticus. Notably, our analyses reveal that *Ca.* Nha. antarcticus selectively takes up a specific set of lipids from its host. Moreover, in co-cultures, the lipidome composition undergoes changes that are likely compensating for an elevated metabolic load as well as enhanced mechanical stress on the host membrane.

Results

Molecular network of lipidome in the *Hrr. lacusprofundi*-*Ca.* Nha. antarcticus system

To determine the lipidome composition of *Ca.* Nha. antarcticus and *Hrr. lacusprofundi*, pure *Ca.* Nha. antarcticus cells were harvested from the nanohaloarchaeal enrichment culture[9] CLAC2B and inoculated into pure cultures of *Hrr. lacusprofundi* in mid-exponential phase at a ratio of 1:10 (*Ca.* Nha. antarcticus to *Hrr. lacusprofundi* cells). Biomass of co-cultures and pure *Hrr. lacusprofundi* was harvested at regular timepoints covering mid- to late-exponential phase of culture growth, whilst pure *Ca.* Nha. antarcticus cells were harvested pre- and post- co-culture growth experiments (Fig. 1a, details in the Method section). To account for potential variation in lipid profiles between *Hrr. lacusprofundi* strains and the possibility for acquisition of lipids from the *Natrinema* sp. present in the enrichment, biomass from 5 additional isolated *Hrr. lacusprofundi* strains and an isolate of the *Natrinema* sp. were used as quality controls (QCs), harvested at mid-exponential phase and subjected to lipidomics. Additionally, to assess whether differences in the *Ca.* Nha. antarcticus lipid profile were due to differences in lipid abundances within the enrichment culture, biomass from the enrichment culture was harvested for lipidomic analysis.

Growth of cultures was assessed by optical density, qPCR measurements, and 16S rRNA-targeted FISH microscopy (Fig. 1b and c, Supplementary Figs. 2 – 5). Optical density readings indicated active growth of both pure *Hrr. lacusprofundi* and co-cultures with the latter exhibiting a slightly slower rate of increase in density than pure culture. The qPCR data showed growth of *Ca.* Nha. antarcticus with an initial doubling of 16S rRNA copy number in the first 12 h followed by an approximately 100-fold increase between 24 and 48 h. Despite lower optical density readings, co-cultures displayed slightly higher 16S rRNA copy numbers for *Hrr. lacusprofundi*. *Hrr. lacusprofundi* 16S rRNA copy number remained stable in both conditions across the 48 h incubation consistent with previously reported genome copy number dynamics reported for other *Halobacteriales*[27]. FISH microscopy revealed multiple stages of interaction between *Ca.* Nha. antarcticus and *Hrr. lacusprofundi* in co-cultures (Supplementary Discussion, Supplementary Figs. 2 - 4). *Hrr. lacusprofundi* cells displayed statistically significant shifts in size across pure and co-cultures with a similar trend in reduction in average cell size over time in both cultures (Supplementary Fig. 5, Supplementary Table 1). Its cell shape also varied significantly but trends were different between cultivation conditions, i.e. pure cultures displayed a gradual trend towards increased circularity, whilst co-cultures displayed a shift towards more elongated rod-shaped cells between 12 and 24 h before increasing in circularity at 48 h though to a lesser degree than in pure cultures.

We analyzed the lipidomes of the samples obtained in our experimental approach with ultra high-pressure liquid chromatography coupled with high-resolution tandem mass spectrometry (UHPLC-HRMS^2) and handled the data obtained with a recently established pipeline that provides a comprehensive analysis of the microbial lipidome in both complex environmental samples and laboratory cultures[28, 29]. In total, 2533 distinct ion components with associated MS^2 spectra were extracted and employed to build up a molecular network (Supplementary Fig. 6). Within this dataset, 1773 ion components (70%) occurred in structure-similarity groupings in the molecular network, while 760 ion components (30%) existed as singletons (i.e., lacking structurally related counterparts). The MS^2 spectra of these ion components did not retrieve any match to any related archaeal lipids in the GNPS spectral library. Lipidome annotation remains a challenge in lipidomic studies, as public spectral databases are inadequately populated. Nevertheless, by comparison with literature[30, 31, 32, 33, 34, 35, 36, 37], as well as tentative identification, we were able to annotate 246 likely archaeal lipids (Supplementary Fig. 6).

These identified 246 lipids can be classified into two primary groups (Fig. 1d). The first group consists of bilayer-forming glycerolipids based on (extended) archaeol (AR, consisting of two ether-bound isoprenoid chains with either 20 or 25 carbon atoms). These occur either as AR core lipids (8 species) or with a polar head group. The latter comprise phospholipids like phosphatidylglycerosulfates (PGS; 13 species), phosphatidylglycerols (PG; 25 species), phosphatidic acids (PA, 3 species), and dimeric phospholipids such as phosphatidylglycerophosphate methyl ester (PGP-Me, 22 species), biphosphatidylglycerols (PGPG, 11 species), and cardiolipins (CL, 41 species), as well as non-phospholipids, including sulphated diglycosyl (SDG, 3 species), monoglycosyl (1G, 2 species), diglycosyl (2G; 2 species) archaeol. The second group comprises non-bilayer forming lipids like menaquinones (MK, 39 species), squalene, and bacterioruberins (15 species).

Additionally, we found 60 unknown species associated with various major archaeal lipid classes. Due to the limited information on MS^2 fragmentation, we were unable to deduce their complete chemical structures. Nonetheless, critical fragments indicated their affiliation with archaeal lipids. For instance, within the 1G/2G/S subnetwork, two unknown lipids were linked to 1G-AR and 2G-AR (Source Data Table), featuring parent ions with *m/z* of 1324.385 and 1306.355. In their MS^2 spectra, both lipids exhibited diagnostic fragment ions at *m/z* 653.681 and 373.368, corresponding to AR, with an assigned elemental composition (AEC) of $C_{43}H_{89}O_3^+$ and lysoAR glycerol with AEC of $C_{23}H_{49}O_3^+$, respectively. Although their exact structures remained elusive, these lipids were included in the overall statistical analysis.

Structural diversity and specificity of lipidome in the *Hrr. lacusprofundi-Ca.* Nha. antarcticus system

To assess the similarity in lipidome composition between the host *Hrr. lacusprofundi*, the symbiont *Ca.* Nha. antarcticus, their co-cultures, and corresponding QCs across multiple time series and replicates, we performed a Principal Component Analysis (PCA) on the abundances of lipid species (Fig. 1e). The first two principal components (PC1 and PC2) accounted for 45.3% of the total lipid variance. The majority of the pure host *Hrr. lacusprofundi* cultures harvested at different time points were closely clustered together, adjacent to the enrichment culture, indicating their high similarity. In contrast, the co-cultures formed a distinct cluster by scoring positively on PC1. The pure *Natrinema* sp. which is the third species isolated from the enrichment and five additional *Hrr. lacusprofundi* strains QC exhibited proximity to the host (depicted in Fig. 1a and Methods section), underscoring the robustness of the culturing experiment and lipidome analysis results. Importantly, a clear separation of the lipidome of *Ca.* Nha. antarcticus, scoring negatively of PC2, from all other samples was observed.

To access the lipidome plasticity of the *Hrr. lacusprofundi-Ca.* Nha. antarcticus system, we also employed an information theory framework[38, 39], quantifying lipidome diversity (H_j index) and specialization (δ_j index) based on the Shannon entropy of the lipidomic frequency distribution. Upon plotting the lipidome specialization and diversity of the *Hrr. lacusprofundi-Ca.* Nha. antarcticus system and QCs, we observed that the pure host *Hrr. lacusprofundi* and QC samples exhibited relatively high lipid diversity and specialization compared to the symbiont (Fig. 1f). In contrast, the co-cultures showed relatively low lipid diversity during the first 12 h of cultivation, which subsequently increased from 24 to 48 h, resulting in a more diverse lipidome profile which was comparable to the profiles of the pure *Hrr. Lacusprofundi* harvested at different time points. The enrichment cultures displayed the highest lipid

specialization, while the symbiont exhibited the lowest lipid diversity and specialization, which may be the result of the symbiont's lack of identifiable unique genes responsible for lipid biosynthesis[9].

We further employed a hierarchical clustering heatmap to examine the distribution of approximately thirty major lipid classes in the host *Hrr. lacusprofundi*, the symbiont *Ca.* Nha. antarcticus, and their co-cultures (Fig. 1g). These lipid classes were categorized based on their polar head groups, degrees of unsaturation, and chain lengths. This method allowed to evaluate the resemblance in lipid class profiles across the cultures and provided an overall comparison of culture similarity grounded in these lipid classifications. The major lipid classes were grouped into six distinct clusters. The first cluster consisted of AR core lipids, MK with a high degree of unsaturation [e.g., MK(8:8) and MK(8:9), the number of the isoprenoid unit in the side chain: the number of double bonds], PGPG and CL with varying degrees of unsaturation. The second cluster comprised phospholipids PG and PGP-Me. The third cluster included PG unknowns, PG-LysoAR (which featured only one C_{20} or C_{25} chain) and bacterioruberins. The remaining clusters comprised highly unsaturated phospholipids (e.g., CL with 10-13 double bonds), 1G, 2G, sulfur-containing lipids, along with several less abundant lipids and unknowns. The lipid class composition of the host sampled at different harvest times were similar to each other, the same as that of the co-cultures obtained at different harvest times. Consistent with the PCA analysis and characterization of the lipidome diversity and specialization, which are both based on the complete lipid species profile, the hierarchical clustering based on the major lipid class compositions also revealed that the host, co-cultures and the symbiont differed from each other.

Changes in the molecular network of the lipidome over time. We next analyzed the variation in the lipidome compositions at the species level by integrating molecular subnetworks of various lipid classes (Fig. 2) from the overall molecular networks (Supplementary Fig. 6), based on their relative abundances. For some lipid species, abundances displayed significant variation between both cultivation conditions and time points of the same cultivation condition (Fig. 3a). The relative abundance of PGP-Me-AR comprised 44% of total lipids in the symbiont, significantly higher ($P < 0.05$, Tukey's Honest Significance Difference test, Fig. 2e) than in the pure culture of the host (13 – 21%) and the co-culture (12 – 31%). In addition, bacterioruberins were underrepresented in symbiont biomass (0.2%) compared to both pure culture of the host (1.2 – 2.9%) and co-culture (1.1 – 5.2%) total lipids, irrespective of time. The pure *Hrr. lacusprofundi* culture also displayed a significantly different lipid profiles compared to that of the co-culture with statistically significant differences in the relative abundance of bacterioruberins, CL, PGP-Me, PG, and MKs of which only the relative abundance of PG was time-dependent (18.5 – 22.3% co-cultures, 7.5 – 17.1% pure *Hrr. lacusprofundi*). The remaining lipid classes displayed time-dependent differences in relative abundance, whereby CL was underrepresented in early-mid co-culture growth (6 – 24 h, 0.6 – 1.2% co-culture, 1.5 – 1.7% pure *Hrr. Lacusprofundi* culture), PGP-Me was overrepresented in early – mid culture growth (0 – 24 h, 20.3 – 31.4% co-culture abundance, 13 – 21% pure *Hrr. lacusprofundi* abundance). In contrast, MKs (12.6 – 12.7% co-culture, 7.7 – 9.2% pure *Hrr. lacusprofundi*) and bacterioruberins (3.6 – 3.7% co-culture, 2.1 – 2.4% pure *Hrr. lacusprofundi*) both became significantly enriched in the late phases (24 – 48 h) of the co-culture biomass compared to the pure culture of *Hrr. lacusprofundi*. There were also statistically significant shifts in the degree of lipid saturation with both purified *Ca.* Nha. antarcticus cells and co-cultures displaying higher rates of saturation in bilayer forming glycerolipids when

compared to pure *Hrr. lacusprofundi* cultures (Supplementary Fig. 7). The degree of saturation of MK also displayed significant variation across samples with *Ca.* Nha. antarcticus and co-cultures in the late growth phase showing increased abundance of MK with one less saturation than the number of the isoprenoid unit in the side chain [e.g., MK(8:7), Supplementary Fig. 7].

Lipid intersection in the *Hrr. lacusprofundi-Ca.* Nha. antarcticus system. We next generated an UpSet plot[40, 41] (Fig. 3b) to illustrate the interrelationship among study objects at the level of the number of lipid species. For instance, it enables to see the differentiation of lipid species that are common to both pure *Hrr. lacusprofundi* and co-cultures, as well as those unique to either. Furthermore, this analysis provides insight into the number of lipid species potentially acquired by the symbiont from the host. As shown in Fig. 3b, the symbiont *Ca.* Nha. antarcticus contained 110 individual lipid species, while the pure culture of *Hrr. lacusprofundi* consistently contained approximately 165 lipid species. The co-cultures had around 140 lipid species in the first 12 h, progressively increasing to approximately 175 species by the end of the experiment at 48 h. This trend aligns with the observed lipid diversity $H_J$index in Fig. 1f, suggesting that *Ca.* Nha. antarcticus incorporated only a limited number of lipid species from *Hrr. lacusprofundi*. Moreover, 86 lipid species across all major lipid classes were observed to be commonly present in all study objects. There were 20 lipid species exclusively found in both co-cultures and pure *Hrr. lacusprofundi* but absent in *Ca.* Nha. antarcticus. Notably, ten specific lipid species, were detected exclusively in the co-cultures between 24 and 40 h. These include two bacterioruberins, manaquinones with equal or additional unsaturation relative to their isoprenoid unit chain length [MK(n:n), MK(n:n+1)], and specific phospholipids. Conversely, three manaquinones with one unsaturation less than their number of isoprenoid units [MK(n:n-1)] were uniquely present in *Ca.* Nha. antarcticus and the enrichment (Fig. 3b and Supplementary Fig. S8), specifically MK(7:6), MK(4:3), and demethylmenaquinone DMK(7:6). This underscores a notable distinction in lipid species composition among the host, co-cultures, and *Ca.* Nha. antarcticus.

Discussion

Given the lack of key lipid biosynthesis pathways in the *Ca*. Nha. antarcticus genome (as shown in Fig. 4), the difference in lipidome composition observed between *Hrr. lacusprofundi* and *Ca*. Nha. antarcticus indicates that *Ca*. Nha. antarcticus selectively acquires specific lipid species from its host. These observations match with the lipid uptake behavior noted for archaeal viruses[42, 43, 44, 45]. For instance, the *Sulfolobus* filamentous virus 1 selectively acquired lipids from its host *Sulfolobus shibatae* for survial[44]. However, our results contrast those of previous studies on two DPANN symbiont-host systems, which did not identify differences between symbiont and host lipid profiles[21, 25]. These discrepancies could be attributed to natural differences between these distantly related symbiotic partnerships or could stem from differences in the methodologies. Specifically, our untargeted lipidome approach provides higher resolution, enhancing the capacity to discern and compare lipid profiles, surpassing traditional methods. It is possible that the application of this technique to other DPANN symbiont host systems may reveal a similar specificity in lipid uptake by those DPANN species. In terms of natural differences, unlike the other reported host archaea which

primarily consist of monolayer membrane tetraether lipids, mostly glycerol dialkyl glycerol tetraethers (GDGTs)[46], the cellular membranes of *Hrr. lacusprofundi* are exclusively formed by bilayer AR lipids[30, 31, 37]. The composition of bilayer-forming intact polar lipids in *Halobacteriales* stands as one of the most extreme instance of negatively charged membranes across the tree of life and is considered to be an adaptation to the high cationic environment[31]. For instance, *Halobacteriales* are the only group of archaeal organisms known for their unique capability to produce both PGP-Me and CLs[33, 35, 47, 48, 49]. This preference for bilayer membranes within *Halobacteriales* confers ecological advantages and the energy-efficient bilayer membrane structure observed in the host, *Hrr. lacusprofundi*, could potentially offer greater flexibility for *Ca*. Nha. antarcticus to selectively uptake specific lipids.

It has previously been shown that the distance between negative charges on the head group structure of CL reduces the efficiency of association with divalent cations such as Mg^{2+} and favors association with monovalent cations such as K^{+}, whilst PGP-Me more efficiently associates with divalent cations[31]. The medium used for these cultivation experiments contains much higher concentrations of Mg^{2+} compared to K^{+} (514 mM combined $MgSO_4$ and MgCl, 39 mM KCl). In addition, experiments in reconstituted phosphatidylcholine based membranes have shown that increased CL abundance reduces both membrane stability and the force necessary for membrane piercing[50]. Given that *Ca.* Nha. antarcticus lacks the capacity to regulate the composition of its membrane; it is plausible that an increased abundance of PGP-Me and a reduced abundance of CL, compared to its host, provides greater stability to the nanohaloarchaeal membrane under such high divalent cation concentrations, reducing the energy expenditure necessary for membrane maintenance.

The significant decrease in bacterioruberin abundance within *Ca.* Nha. antarcticus biomass compared to *Hrr. lacusprofundi* suggests a counter-selection against recruitment of bacterioruberins into the symbiont membrane. Bacterioruberin functions as an antioxidant, increases membrane rigidity, and is associated with rhodopsins within membranes of members of the *Halobacteriales*[51]. The exclusion of bacterioruberins from the nanohaloarchaeal membrane may reflect lower oxidative stress due to the fermentative lifestyle predicted from the genome or pressure to maintain higher membrane fluidity. In addition, whilst *Ca.* Nha. antarcticus possesses a rhodopsin gene, this is predicted to encode a sensory rhodopsin, in contrast to the rhodopsins that are present in the *Hrr. lacusprofundi* genome that likely generate proton gradients for ATP production and may necessitate more bacterioruberins. Consistent with oxidative stress playing a role in preference for membrane composition, *Ca.* Nha. antarcticus also displayed a preference for MKs with increased degree of unsaturation, which has been suggested to operate more efficiently in hypoxic conditions[50]. Interestingly, for *Mycobacterium tuberculosis* this was proposed as an adaptation to intracellular environments, which are often lower in oxygen content[50]. Recently it was reported that *Ca.* Nha. antarcticus appears to invade host cells during the process of interaction[26] and the preference for increased MK desaturation may similarly assist in survival during this stage of the symbiont's lifecycle. The contrasting metabolic strategies employed by this nanohaloarchaeum compared to its host may therefore favor the selective acquisition of lipids seen in our data to maximize metabolic efficiency and survival.

In addition to the differences between *Ca.* Nha antarcticus and host lipid profiles, there were significant shifts in lipidome composition of the co-culture biomass and pure *Hrr. lacusprodundi*. This indicates a different membrane

composition for host cells during co-culture. Initially co-cultures showed low lipid species numbers and lipid diversity during the first 12 h, which may partly be due to the limited lipid diversity of the symbiont *Ca*. Nha. antarcticus but also reflects changing lipid composition in the host *Hrr*. *lacusprofundi*. The subsequent rise in the number of lipid species and diversity in co-cultures over the following 36 h, along with the distinct lipid composition compared to the pure *Hrr*. *lacusprofundi*, suggests that *Hrr*. *lacusprofundi* modifies its membrane composition in response to interactions with *Ca.* Nha. antarcticus. Interestingly, lipid species enriched in *Ca.* Nha. antarcticus (e.g. PGP-Me) showed a decrease in abundance over the course of the time series in the co-cultures. In contrast, the diversity and abundance of other lipids, such as bacterioruberins and MKs, which were less prevalent in the nanohaloarchaeal membrane, showed an increase in abundance in the co-cultures. The increased abundance of bacterioruberins and MKs likely indicates a response from *Hrr*. *lacusprofundi* to an increase in metabolic load due to the presence of the symbiont. Similarly, bilayer-forming glycerolipids showed an increase in the rate of desaturation resulting in increased membrane fluidity and enhanced permeability for electron transport and therefore more efficient respiration[31, 52].

Similar to *Ca.* Nha antarcticus, the CL abundance in the co-culture biomass also decreased, which likely reflects decreased abundance in both species' bilayers. As mentioned above, the reduced CL abundance within the membrane may act to increase membrane stability and resistance to mechanical stress. DPANN archaea that take up nutrients directly from host cells require access to their host's cytoplasm for nutrient acquisition and have been observed to form channels in host membranes as part of interactions[53]. Therefore, it seems possible that the decreased CL content in the *Hrr*. *lacusprofundi* membrane fortifies the membrane, potentially as a defense mechanism to prevent infection by its symbiont or to enhance survival of infected cells by reducing the chance of membrane destabilization.

Our study revealed that the DPANN archaeon *Ca*. Nha. antarcticus selectively acquires specific lipids from its host *Hrr. lacusprofundi.* Additionally, during co-cultivation, *Hrr. lacusprofundi* modified its own lipid composition likely resulting in changes in membrane integrity that may constitute a lipid defense mechanism to restrict the exploitability of host cells by the symbiont. This study also emphasizes the strength of employing computational untargeted lipidomics approaches to elucidate lipid interactions within a host-symbiont culture system. Further research is needed to delve into the mechanisms underlying the specific selection of lipids and lipid defensive functions by the symbiont and host, respectively. Understanding these mechanisms will deepen our knowledge of archaeal host-symbiont interactions at a molecular level, especially in the context of lipid exchange and survival strategies in extreme environments.

Methods

Purification of Nanohaloarchaeal Cells

To acquire *Ca.* Nha. antarcticus cells for use in co-culture experiments, nanohaloarchaeal cells were purified via filtration following previously described methods[26]. Briefly, 1 L of nanohaloarchaeal enrichment culture[9] was filtered sequentially through 0.8 µm (3x) and 0.2 µm (3x) polycarbonate filters (Isopore, Merck Scientific). Purified cells were then pelleted by centrifugation at 20,000 *g* for 10 min and resuspended in 4 mL of DBCM2 media[54]. 100 µL of purified

cells were stained with MitoTracker Green (as previously described[26]) and Nile Red (1 µg/mL in 30% Salt Water (SW) mix[54]) for 1 h and imaged on an Axio Imager M2 microscope to assess purity and possible contamination of cell debris. 100 uL of purified cells were pelleted, DNA extracted using a PeqGold DNA Blood and Tissue Extraction kit following the manufacturer's instructions (VWR), and PCR performed targeting both *Ca.* Nha. antarcticus and *Hrr. lacusprofundi* (primer details in Table 1) to confirm purity of cells. Remaining *Ca.* Nha. antarcticus cells were divided into four aliquots (~$5x10^7$ cells): one aliquot was pelleted and used as biomass for lipidomic analyses, whilst the other three aliquots were inoculated into pure cultures of *Hrr. lacusprofundi* R1S1 for co-culture experiments.

Cultivation Experiments

Cultures of *Hrr. lacusprofundi* R1S1B were grown in DBCM2 in triplicate by shaking (120 r.p.m.) at 30 °C in volumes of 250 mL to late exponential phase. Cultures were then split into duplicates and diluted to an OD_{600} of 0.2 in 250 mL. Infected cultures were inoculated with ~$5x10^7$ purified nanohaloarchaeal cells (see above). Samples for downstream analyses were harvested at 0 h, 6 h, 12 h, 24 h, and 48 h. Culture density was measured using OD_{600} in triplicate with fresh DBCM2 media as blank at each timepoint. For lipidomics samples, 10 mL of culture was pelleted at 6,000 *g* for 30 min then washed in 30% SW and pelleted at 20,000 *g* for 10 min three times to remove excess media components. Samples for qPCR were acquired by pelleting 1 mL of culture at 20,000 *g*, and removal of supernatant, and DNA extraction were performed as above. Samples for FISH were fixed with 2.5% glutaraldehyde at 4 °C overnight, washed with milliQ water and stored at -20 °C. In addition to *Hrr. lacusprofundi* R1S1 cultures and *Hrr. lacusprofundi* R1S1B – *Ca.* Nha. antarcticus co-cultures, lipidomics samples were collected from the nanohaloarchaeal enrichment culture, a pure *Natrinema* sp. isolated from the enrichment, and five additional and distinct *Hrr. lacusprofundi* strains as controls.

PCR and qPCR

All PCR and qPCR reactions were performed using the same primer pairs for either *Ca.* Nha. antarcticus or *Hrr. lacusprofundi* (Table 1). Purity of filtered *Ca.* Nha. antarcticus cells was assessed by standard PCR with both sets of primers using DreamTaq polymerase (ThermoFisher) for 35 cycles at 55 °C annealing temperature. Standards for qPCR were produced by PCR amplification of both *Ca.* Nha. antarcticus and *Hrr. lacusprofundi* 16S rRNA genes followed by cloning of amplified products into a pGEM-T easy vector (Promega), transformation of JM109 competent cells (Promega), and subsequent plasmid purification using a peqGOLD Plasmid Miniprep Kit, all steps were carried out as per the manufacturer's instructions. qPCR reactions were performed in a CFX96 Real-Time PCR Detection System (Bio-Rad) for 40 cycles with an annealing temperature of 55 °C.

Fluorescence Microscopy

Fluorescence *in-situ* Hybridisation reactions were carried out as previously described[9, 55]. Briefly, fixed samples were pelleted at 20,000 *g* for 10 min, resuspended in hybridisation buffer with probes (100 pM working concentration, probe details in Table 1) specific to *Ca.* Nha. antarcticus and *Hrr. lacusprofundi* and incubated at 46 °C for 3 h. Cells were then pelleted at 20,000 *g* for 10 min, resuspended in wash buffer, and incubated at 48 °C for 30 min. Cells were then pelleted again at 20,000 g for 10 min, resuspended in PBS with 300 nM DAPI for counterstaining, and incubated

for 1 h. Stained cells (30 µL) were mounted onto glass slides with antifadant and imaged on an Axio Imager M2. Data analysis and image processing was conducted using Fiji[56].

Genomic Analyses

To identify lipid biosynthesis capacities within *Ca.* Nha. antarcticus and *Hrr. lacusprofundi* genomes were annotated with a suite of functional annotation tools and manually curated. Due to rapid genome rearrangement within *Hrr. lacusprofundi* strains *Hrr. lacusprofundi* strain R1S1 was re-sequenced and assembled prior to conducting experiments. To avoid confusion with the existing *Hrr. lacusprofundi* R1S1 genome, we designate this as *Hrr. lacusprofundi* R1S1B. DNA was extracted as described above and sent for sequencing with Illumina 2x 150 bp paired end reads (Eurofins). Raw reads were trimmed with trimmomatic[57] (v0.36, settings: SLIDINGWINDOW:5:22 MINLEN:100), assembly was performed using SPAdes[58] (v3.15.0, settings: -t 10 -k 21,39,59,99,127) and assembly was quality assessed with quast[59] (v4.6.3). Genomic analysis of *Ca.* Nha. antarcticus was carried out on the published genome accessed from IMG (IMG Genome ID: 2643221421). For both genomes the approach from Dombrowski *et. al.* (2020)[6] was followed. Briefly, coding sequences were predicted using Prokka v1.14[60] (settings: --kingdom Archaea --addgenes --force --increment 10 --compliant --centre UU --cpus 20 --norrna –notrna). For functional annotation of genes several additional databases were used including COGs[61] (downloaded October 2020), arCOGs[62] (2014 version), KO profiles from the KEGG Automated Annotation Server[63] (downloaded April 2021), the Pfam database[64] (release 34.0), the TIGRFAM database[65] (release 15.0), the Carbohydrate-Active enZymes (CAZy) database[66] (v7, downloaded August 2020), the Transporter Classification Database[67] (downloaded April 2021), the Hydrogenase database [68] (HydDB, downloaded July 2020), and NCBI_nr (downloaded Aug 2021). In addition to this, protein domain predictions were carried out using InterProScan[69] (v5.62-94.0, setting: --iprlookup --goterms).

Annotations for the respective databases were carried out as follows. COGs, arCOGs, KOs, PFAMs, TIGRFAMs, and CAZymes were all identified using hmmsearch v3.1b2 (settings: -E 1e-5). The Transporter Classification Database and Hydrogenase Database were queried using BLASTp[70] v2.7.1 (settings: -evalue 1e-20). For database searches the best hit was selected based on highest e-value and bit-score and summarized in Supplementary Tables 2 and 4. Multiple hits were allowed for InterProScan domain annotations using a custom script for parsing results (parse_IPRdomains_vs2_GO_2.py). Best blast hits against the NCBI_nr database were identified using DIAMOND[71] (settings: blastp --more-sensitive --evalue 1e-5 --no-self-hits). Identification of lipid biosynthesis genes was performed manually by screening annotated genes for KEGG annotations associated with synthesis of relevant lipids.

Lipidome extraction and analysis

The methodology for lipid extraction and measurement in this study is detailed in Bale et al.[29]. To eliminate background lipids and contaminants, both medium blanks and extraction blanks were utilized. Briefly, the samples and blanks were extracted using a modified Bligh-Dyer extraction method[72, 73]. They were subjected to ultrasonic extraction for 10 min, twice using a mixture of methanol, dichloromethane (DCM) and [PO_4^{3-}] buffer (2:1:0.8, v/v/v)

and twice with a mixture of methanol, DCM and aqueous trichloroacetic acid solution at pH 3 in the same ratio. The organic phase was separated by adjusting the solvent mixture to a final ratio of 1:1:0.9 (v/v/v) with additional DCM and buffer. This organic phase was then subjected to three further extractions using DCM and then dried under a stream of N_2 gas. The dry extract was re-dissolved in a methanol and DCM mixture (9:1, v/v), followed by filtration through 0.45 μm regenerated cellulose syringe filters (4 mm diameter; Grace Alltech). The filtered extracts were subsequently analyzed using an Agilent 1290 Infinity I UHPLC system coupled to a Q Exactive Orbitrap MS (Thermo Fisher Scientific, Waltham, MA). The generated output data from the UHPLC-HRMS^2 analysis were processed with MZmine software[74] to extract MS^1 and MS^2 spectra as well as quantify peaks. This processing included several steps: mass peak detection, chromatogram building, deconvolution, isotope grouping, feature alignment, and gap filling (https://ccms-ucsd.github.io/GNPSDocumentation).

Molecular Networking

The MS/MS spectra dataset was further processed using the Feature-Based Molecular Networking tool on the Global Natural Product Social Molecular Networking (GNPS) platform[75, 76]. Molecular networking a key data analysis methodology in untargeted metabolomics studies based on MS/MS analysis, arranges MS/MS spectra into a network-like map. In this map, molecules with similar spectral patterns are clustered together, indicating their structural similarities. The analysis involves calculating vector similarities comparing pairs of spectra based on at least five matching fragment ions (peaks). This comparison not only considers the relative intensities of the fragment ions but also the difference in precursor *m/z* values between the spectra[76, 77]. The molecular network is constructed using MATLAB scripts, where each spectrum is linked to its top K scoring matches, usually allowing up to 10 connections per node. Connections (edges) between spectra are retained if they rank among the top K matches for both spectra and if the vector similarity score surpasses a predetermined threshold. The similarity score is quantified as a cosine value, where a score of 1.0 signifies identical spectra. In this study, a cosine value of 0.5 was used to define significant spectral similarities, indicating a moderate to high level of structural resemblance between the analyzed molecules.

In the molecular networking analysis of the MS/MS spectra, when an ion component displayed both protonated $[M+H]^+$ and ammoniated $[M+NH_4]^+$ ions, the overall abundance of that component was calculated as the combined total of the abundances of these two ion forms. For the construction of the molecular network, a minimum of five shared fragment ions was established as the criterion for connecting pairs of related MS/MS spectra with an edge. Each node within the network was permitted to connect to a maximum of ten analogs. In addition, consensus spectra were compared against the GNPS spectral library[76, 78], allowing for a maximum analog mass difference of *m/z* 500. The maximum size of nodes allowed in a single connected subnetwork was capped at 100. In scenarios where the dataset contained a significant number of related lipids (exceeding 100), these lipids were segregated into different subnetworks.

The molecular networks derived from the analysis were visualized using Cytoscape version 3.9.1[79, 80]. It is important to note that since many of the lipids detected in this study have not been previously characterized, authentic standards for absolute quantification were not available. The lipids were corrected for sample recovery with a 1,2-dipalmitoyl-

sn-glycero-3-O-4'-[N,N,N-trimethyl(d9)]-homoserine (DGTS-d9) internal standard then examined based on their calibrated peak area responses. Consequently, the relative peak areas calculated abundance do not necessarily indicate the actual relative abundance of different lipids in samples. Nevertheless, this method allowed comparison of lipids between different cultures or cultivation conditions, rather than determining the absolute quantities of each lipid present [81].

Information Theory framework

The lipidome's diversity and specialization, along with the specificity of individual lipid species, were defined and analyzed using an information theory framework[38, 39, 82]. Lipids were characterized via their distinct tandem MS^2 spectra and their relative occurrence frequencies across various cultures. The lipidome diversity, the H_j index, was calculated using the Shannon entropy based on the frequency distribution of lipid species as determined by the abundance of their MS^2 precursor ions. The equation is as follows

$$H_j = -\sum_{i=1}^{m} P_{ij} log_2(P_{ij})$$

where P_{ij} correspond to the relative frequency of the *i*th MS^2 (i = 1, 2, …, m) in the *j*th sample (j = 1, 2, …, t), to illustrate how abundant a specific MS^2 spectrum is relative to all others.

The average frequency of the *i*th MS^2 among samples was calculated as

$$P_i = \frac{1}{t}\sum_{j=1}^{m} P_{ij}$$

Individual lipid species specificity, the S_i index, was defined as the identity of a given MS^2 regarding frequencies among all the cultures. The lipid species specificity was calculated as

$$S_i = \frac{1}{t}\left(\sum_{j=1}^{t} \frac{P_{ij}}{P_i} log_2 \frac{P_{ij}}{P_i}\right)$$

Individual lipid species specificity of specific cultures, was defined as S_{ij} index.

$$S_{ij} = \sum_{j=1}^{t} \frac{P_{ij}}{P_i} log_2 \frac{P_{ij}}{P_i}$$

The lipidome specialization δ_j index was measured as the average of the MS^2 specificities using the following formula

$$\delta_j = \sum_{i=1}^{m} P_{ij}\, S_i$$

Statistical analysis

For principal component analysis (PCA), the abundance data of lipid species were initially transformed using the Hellinger distance method[83] to mitigate bias arising from zero values. This data was then processed and visualized

using R software, version 4.1.2. Hierarchical clustering was performed using the "ggplot2" and "pheatmap" packages in R, version 4.3.2.

Figure captions

Figure 1. Overview of the experimental design and the general lipidome composition in the *Hrr. lacusprofundi*-*Ca.* Nha. antarcticus system. (a) To investigate the lipidomes of pure *Hrr. lacusprofundi*, co-cultures of *Hrr. lacusprofundi* with *Ca.* Nha. antarcticus, and pure *Ca.* Nha. antarcticus, we performed 48 h long incubations with sampling at regular time points. The pure host culture and co-culture were sampled at 0, 6, 12 ,24, and 48 h and collected for growth measurements (Optical density, qPCR), visual analysis (FISH), and lipidomics. *Ca.* Nha. antarcticus cells were purified from the enrichment culture at the beginning of the experiment and from co-cultures at 48 h for lipidomic analysis. (**b**) qPCR-based growth measurements of pure *Hrr. lacusprofundi* cultures and co-cultures of *Hrr. lacusprofundi* with *Ca.* Nha. antarcticus. Error bars show the standard deviation of calculated 16S rRNA gene copy number. (**c**) Optical density at 600 nm (OD_{600}) growth measurements of pure *Hrr. lacusprofundi* cultures and co-cultures of *Hrr. lacusprofundi* with *Ca.* Nha. antarcticus. Error bars show the standard deviation of measured OD_{600} values. (**d**) The number of individual lipid species in major lipid classes among all the samples. (**e**) Principal Component Analysis (PCA) based on the abundance of intact polar lipid species, showcasing the variance in general lipidomic features among distinct cultures or over varying culture durations. (**f**) Information theory analysis showing lipidome diversity (H_j index) and specialization (δ_j index) based on the Shannon entropy of the lipidomic frequency distribution. Error bars in the data represent variability across replicates. (**g**) Hierarchical clustering heatmap depicting the distribution of major lipid classes among distinct cultures or over varying culture durations. The colour bar on the right side represents Z-score normalization scale (ranges from -3 to +3 standard deviation). Sample abbreviations: *Ca*. Nha. antarcticus (Nha), *Hrr. lacusprofundi* (HP), co-cultures (Cc). Lipid abbreviations: archaeol core lipids (AR), phosphatidylglycerol (PG), phosphatidylglycerosulfate (PGS), phosphatidic acid (PA), phosphatidylglycerophosphate methyl ester (PGP-Me), biphosphatidylglycerol (PGPG), cardiolipin (CL), sulphated diglycosyl (SDG), monoglycosyl (1G), diglycosyl (2G), archeaol lipids containing a sulfur-containing head group except for PGS (S), menaquinone (MK), an “extended “ archaeological chain”, i.e. with a C_{25} isoprenoid carbon chain (EXT-AR), unsaturation in the archaeol chain (uns). The two “n” in MK (n:n) stand for numbers of the isoprenoid unit in the side chain and unsaturation in the isoprenoid chain, respectively. MK(n:n-1) signifies one less double bond in the nth isoprenoid chain.

Figure 2. Time-dependent changes in the relative abundances of individual lipids indicated in a combined molecular network of lipidomes of the *Hrr. lacusprofundi*-*Ca.* Nha. antarcticus system. (a) *Ca.* Nha. antarcticus isolated from the co-culture after 48 h of growth, (**b**) a pure culture of *Hrr. lacusprofundi* harvested at five different time points, 0, 6, 12, 24 and 48 h. (**c**) a co-culture of Ca. Nha. antarcticus and its host *Hrr. lacusprofundi* harvested at 0, 6, 12, 24 and 48 h of growth. The nodes in the molecular network represent MS/MS spectra of ion components (lipids) which are connected based on spectral similarity. The sizes of the lipid nodes indicate their average relative

abundance. General structures are shown in (**a**). The abbreviations of lipids are given in the caption of Fig. 1. The network presented here was a composite derived from merging the lipid subnetworks shown in Supplementary Fig. S6.

Figure 3. The presence, absence, and changes in lipid composition in the *Hrr. lacusprofundi*-*Ca.* Nha. antarcticus system (a) The relative abundance of representative lipid species within the most dominant lipid classes. Statistical differences in lipid species among the samples were assessed using the Tukey's Honest Significance Difference test (TukeyHSD), and results were visualized with the Compact Letter Display (CLD) ($P < 0.05$). (**b**) The intersection of lipid species across samples is illustrated through an UpSet plot[40, 41]. A threshold of 0.01% relative abundance of total lipids was applied to determine the presence of a lipid in a specific sample; lipids with less than 0.01% of total lipid abundance were considered absent in that sample. The dark connected dots denote lipid species shared among these samples. Abbreviations: demethylmenaquinone (DMK)[84], methylmenaquinone (MMK), dimethylmenaquinone (DMMK). The representative lipid species are 1G-AR (*m/z* 832.760, $C_{49}H_{102}O_8N^+$), 2G-AR (*m/z* 994.813, $C_{55}H_{112}O_{13}N^+$), AR (*m/z* 653.681, $C_{43}H_{89}O_3^+$), Bacterioruberin (*m/z* 741.581, $C_{50}H_{77}O_4^+$), CL-AR-AR (*m/z* 1522.313, $C_{89}H_{183}O_{13}P_2^+$), MK(8:8) (*m/z* 717.560, $C_{51}H_{73}O_2^+$), PG-AR (*m/z* 807.684, $C_{46}H_{96}O_8P^+$), PGP-Me-AR (*m/z* 901.666, $C_{47}H_{99}O_{11}P_2^+$), PG-PG-AR (*m/z* 961.687, $C_{49}H_{103}O_{13}P_2^+$).

Figure 4. A schematic figure showing lipid composition and biosynthetic pathways in *Hrr. lacusprofundi, Ca.* Nha. antarcticus, and co-cultures. Lipid biosynthetic pathways were manually reconstructed using genome annotations of *Hrr. lacusprofundi* and *Ca.* Nha. antarcticus inferred using KEGG orthology (List of genes present in Supplementary Tables 3-5). In some cases (marked by *) the correct KEGG annotation could not be identified but a related enzyme that may carry out the reaction was present and is shown instead. Insets above cells show abundance of different lipid classes in each condition at 24 h. Localisation of MKs within the bilayer is shown in two possible states in line with current uncertainty regarding the exact position MKs occupies within the membrane[31]. A representative 16s rRNA target FISH microscopy image as an example of the interactions between host and symbiont is included (Colours correspond to DNA: Blue, *Hrr. lacusprofundi*: Yellow, *Ca.* Nha. antarcticus: Magenta).

Acknowledgments

We thank Michel Koenen for the lipidome extraction. We acknowledge Wencong Huang and Dina Castillo Boukhchtaber for their support in the culturing experiment and Prof. Stefan Schouten, Prof. Laura Villanueva, Dr. Kerstin Fiege and Dr. Diana Sahonero Canavesi who provided valuable comments on the initial experiment. JSSD received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement no.694569—MICROLIPIDS) and from a Spinoza award from NWO. AS has

received funding from the NWO-I foundation of the Netherlands Organisation for Scientific Research (WISE fellowship to AS), the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 947317, ASymbEL), the Moore–Simons Project on the Origin of the Eukaryotic Cell, Simons Foundation 735929LPI (https://doi.org/10.46714/735929LPI), and a Gordon and Betty Moore Foundation's Symbiosis in Aquatic Systems Initiative (GBMF9346).

Data availability

The processed data (.mgf and .csv) with the molecular network and detailed parameter settings can be accessed at the GNPS platform [https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=c75ddcdd8d2e426e9d537ee1037a2b43]. The *Halorubrum lacusprofundi* R1S1B genome is available at NCBI under BioProject: PRJNA1046704, BioSample: SAMN38507334. The source data for the figures are provided in the extended dataset in our data repository on zenodo [https://doi.org/10.5281/zenodo.10258365].

Code Availability

All custom scripts and workflows used to generate data can be found in our data repository on zenodo [https://doi.org/10.5281/zenodo.10258365].

Author contribution

S.D., J.N.H., N.J.B., J.S.S.D., and A.S. conceived the study. J.N.H. performed all cultivation and microscopy analyses. S.D. performed lipidome data analyses. S.D. and N.J.B. performed lipid identifications. S.D. and J.N.H. wrote the draft manuscript. J.S.S.D. and A.S. supervised the study. All authors read, discussed, and approved the final version of the manuscript.

References

1. Rinke C*, et al.* Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499**, 431-437 (2013).

2. Castelle Cindy J*, et al.* Genomic Expansion of Domain Archaea Highlights Roles for Organisms from New Phyla in Anaerobic Carbon Cycling. *Current Biology* **25**, 690-701 (2015).

3. Probst AJ*, et al.* Differential depth distribution of microbial function and putative symbionts through sediment-hosted aquifers in the deep terrestrial subsurface. *Nature Microbiology* **3**, 328-336 (2018).

4. Baker BJ*, et al.* Enigmatic, ultrasmall, uncultivated Archaea. *Proceedings of the National Academy of Sciences* **107**, 8806-8811 (2010).

5. Probst AJ*, et al.* Biology of a widespread uncultivated archaeon that contributes to carbon fixation in the subsurface. *Nature Communications* **5**, 5497 (2014).

6. Dombrowski N*, et al.* Undinarchaeota illuminate DPANN phylogeny and the impact of gene transfer on archaeal evolution. *Nature Communications* **11**, 3939 (2020).

7. Castelle CJ, Banfield JF. Major New Microbial Groups Expand Diversity and Alter our Understanding of the Tree of Life. *Cell* **172**, 1181-1197 (2018).

8. Vavourakis CD*, et al.* Metagenomic Insights into the Uncultured Diversity and Physiology of Microbes in Four Hypersaline Soda Lake Brines. *Frontiers in Microbiology* **7**, (2016).

9. Hamm JN*, et al.* Unexpected host dependency of Antarctic Nanohaloarchaeota. *Proceedings of the National Academy of Sciences* **116**, 14661-14670 (2019).

10. Castelle CJ, Brown CT, Anantharaman K, Probst AJ, Huang RH, Banfield JF. Biosynthetic capacity, metabolic variety and unusual biology in the CPR and DPANN radiations. *Nature Reviews Microbiology* **16**, 629-645 (2018).

11. Liu X*, et al.* Insights into the ecology, evolution, and metabolism of the widespread Woesearchaeotal lineages. *Microbiome* **6**, 102 (2018).

12. Durbin AM, Teske A. Sediment-associated microdiversity within the Marine Group I Crenarchaeota. *Environmental Microbiology Reports* **2**, 693-703 (2010).

13. Huber H, Hohn MJ, Rachel R, Fuchs T, Wimmer VC, Stetter KO. A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature* **417**, 63-67 (2002).

14. Jahn U*, et al.* Nanoarchaeum equitans and Ignicoccus hospitalis: New Insights into a Unique, Intimate Association of Two Archaea. *Journal of Bacteriology* **190**, 1743-1750 (2008).

15. Dombrowski N, Lee J-H, Williams TA, Offre P, Spang A. Genomic diversity, lifestyles and evolutionary origins of DPANN archaea. *FEMS Microbiology Letters* **366**, fnz008 (2019).

16. Wurch L*, et al.* Genomics-informed isolation and characterization of a symbiotic Nanoarchaeota system from a terrestrial geothermal environment. *Nature Communications* **7**, 12115 (2016).

17. St. John E*, et al.* A new symbiotic nanoarchaeote (Candidatus Nanoclepta minutus) and its host (Zestosphaera tikiterensis gen. nov., sp. nov.) from a New Zealand hot spring. *Systematic and Applied Microbiology* **42**, 94-106 (2019).

525
526 18. Golyshina OV*, et al.* 'ARMAN' archaea depend on association with euryarchaeal host in culture
527 and in situ. *Nature Communications* **8**, 60 (2017).

528
529 19. Chen L-X*, et al.* Metabolic versatility of small archaea Micrarchaeota and Parvarchaeota. *The ISME*
530 *Journal* **12**, 756-775 (2018).

531
532 20. La Cono V*, et al.* Symbiosis between nanohaloarchaeon and haloarchaeon is based on utilization
533 of different polysaccharides. *Proceedings of the National Academy of Sciences* **117**, 20223-20234
534 (2020).

535
536 21. Krause S*, et al.* The importance of biofilm formation for cultivation of a Micrarchaeon and its
537 interactions with its Thermoplasmatales host. *Nature Communications* **13**, 1735 (2022).

538
539 22. Giannone RJ*, et al.* Life on the edge: functional genomic response of Ignicoccus hospitalis to the
540 presence of Nanoarchaeum equitans. *The ISME Journal* **9**, 101-114 (2015).

541
542 23. Munson-McGee Jacob H, Field Erin K, Bateson M, Rooney C, Stepanauskas R, Young Mark J.
543 Nanoarchaeota, Their Sulfolobales Host, and Nanoarchaeota Virus Distribution across
544 Yellowstone National Park Hot Springs. *Applied and Environmental Microbiology* **81**, 7860-7868
545 (2015).

546
547 24. Sakai HD*, et al.* Insight into the symbiotic lifestyle of DPANN archaea revealed by cultivation and
548 genome analyses. *Proceedings of the National Academy of Sciences* **119**, e2115449119 (2022).

549
550 25. Jahn U, Summons R, Sturt H, Grosjean E, Huber H. Composition of the lipids of Nanoarchaeum
551 equitans and their origin from its host Ignicoccus sp. strain KIN4/I. *Archives of Microbiology* **182**,
552 404-413 (2004).

553
554 26. Hamm JN*, et al.* The intracellular lifestyle of an archaeal symbiont. *bioRxiv*,
555 2023.2002.2024.529834 (2023).

556
557 27. Breuert S, Allers T, Spohn G, Soppa J. Regulated Polyploidy in Halophilic Archaea. *PLOS ONE* **1**, e92
558 (2006).

559
560 28. Ding S*, et al.* Lipidomics of environmental microbial communities. II: Characterization using
561 molecular networking and information theory. *Frontiers in Microbiology* **12**, (2021).

562
563 29. Bale NJ*, et al.* Lipidomics of environmental microbial communities. I: Visualization of component
564 distributions using untargeted analysis of high-resolution mass spectrometry data. *Frontiers in*
565 *Microbiology* **12**, (2021).

30. Bale NJ*, et al.* New Insights Into the Polar Lipid Composition of Extremely Halo(alkali)philic Euryarchaea From Hypersaline Lakes. *Frontiers in Microbiology* **10**, (2019).

31. Kellermann MY, Yoshinaga MY, Valentine RC, Wörmer L, Valentine DL. Important roles for membrane lipids in haloarchaeal bioenergetics. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1858**, 2940-2956 (2016).

32. Elling FJ*, et al.* Chemotaxonomic characterisation of the thaumarchaeal lipidome. *Environmental Microbiology* **19**, 2681-2700 (2017).

33. Angelini R, Corral P, Lopalco P, Ventosa A, Corcelli A. Novel ether lipid cardiolipins in archaeal membranes of extreme haloalkaliphiles. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1818**, 1365-1373 (2012).

34. Yoshinaga MY, Kellermann MY, Rossel PE, Schubotz F, Lipp JS, Hinrichs K-U. Systematic fragmentation patterns of archaeal intact polar lipids by high-performance liquid chromatography/electrospray ionization ion-trap mass spectrometry. *Rapid Communications in Mass Spectrometry* **25**, 3563-3574 (2011).

35. Koga Y, Morii H. Recent Advances in Structural Research on Ether Lipids from Archaea Including Comparative and Physiological Aspects. *Bioscience, Biotechnology, and Biochemistry* **69**, 2019-2034 (2005).

36. Gibson JAE, Miller MR, Davies NW, Neill GP, Nichols DS, Volkman JK. Unsaturated diether lipids in the psychrotrophic archaeon Halorubrum lacusprofundi. *Systematic and Applied Microbiology* **28**, 19-26 (2005).

37. Yao W*, et al.* Lipidomic chemotaxonomy aligned with phylogeny of Halobacteria. *Frontiers in Microbiology* **14**, (2023).

38. Li D, Halitschke R, Baldwin IT, Gaquerel E. Information theory tests critical predictions of plant defense theory for specialized metabolism. *Science Advances* **6**, eaaz0381 (2020).

39. Martínez O, Reyes-Valdés MH. Defining diversity, specialization, and gene specificity in transcriptomes through information theory. *Proceedings of the National Academy of Sciences* **105**, 9709 (2008).

40. Lex A, Gehlenborg N, Strobelt H, Vuillemot R, Pfister H. UpSet: Visualization of Intersecting Sets. *IEEE Transactions on Visualization and Computer Graphics* **20**, 1983-1992 (2014).

607 41. Conway JR, Lex A, Gehlenborg N. UpSetR: an R package for the visualization of intersecting sets
608 and their properties. *Bioinformatics* **33**, 2938-2940 (2017).

609
610 42. Baquero DP*, et al.* A filamentous archaeal virus is enveloped inside the cell and released through
611 pyramidal portals. *Proceedings of the National Academy of Sciences* **118**, e2105540118 (2021).

612
613 43. Wang F*, et al.* A packing for A-form DNA in an icosahedral virus. *Proceedings of the National*
614 *Academy of Sciences* **116**, 22591-22597 (2019).

615
616 44. Liu Y*, et al.* Structural conservation in a membrane-enveloped filamentous virus infecting a
617 hyperthermophilic acidophile. *Nature Communications* **9**, 3360 (2018).

618
619 45. Kasson P*, et al.* Model for a novel membrane envelope in a filamentous hyperthermophilic virus.
620 *eLife* **6**, e26268 (2017).

621
622 46. Schouten S, Hopmans EC, Sinninghe Damsté JS. The organic geochemistry of glycerol dialkyl
623 glycerol tetraether lipids: A review. *Organic Geochemistry* **54**, 19-61 (2013).

624
625 47. Corcelli A, Lobasso S. 25 Characterization of Lipids of Halophilic Archaea. In: *Methods in*
626 *Microbiology*). Academic Press (2006).

627
628 48. Sprott GD, Dicaire CJ, Gurnani K, Deschatelets LA, Krishnan L. Liposome adjuvants prepared from
629 the total polar lipids of Haloferax volcanii, Planococcus spp. and Bacillus firmus differ in ability to
630 elicit and sustain immune responses. *Vaccine* **22**, 2154-2162 (2004).

631
632 49. van de Vossenberg JLCM, Driessen AJM, Grant D, Konings WN. Lipid membranes from halophilic
633 and alkali-halophilic Archaea have a low H+ and Na+ permeability at high salt concentration.
634 *Extremophiles* **3**, 253-257 (1999).

635
636 50. Upadhyay A*, et al.* Partial Saturation of Menaquinone in Mycobacterium tuberculosis: Function
637 and Essentiality of a Novel Reductase, MenJ. *ACS Central Science* **1**, 292-302 (2015).

638
639 51. Giani M, Garbayo I, Vílchez C, Martínez-Espinosa RM. Haloarchaeal Carotenoids: Healthy Novel
640 Compounds from Extreme Environments. *Mar Drugs* **17**, (2019).

641
642 52. Siliakus MF, van der Oost J, Kengen SWM. Adaptations of archaeal and bacterial membranes to
643 variations in temperature, pH and pressure. *Extremophiles* **21**, 651-670 (2017).

644
645 53. Junglas B*, et al.* Ignicoccus hospitalis and Nanoarchaeum equitans: ultrastructure, cell–cell
646 interaction, and 3D reconstruction from serial sections of freeze-substituted cells and by electron
647 cryotomography. *Archives of Microbiology* **190**, 395-408 (2008).

648
649 54. Dyall-Smith M. The Halohandbook v7.3. (2015).

650
651 55. Pernthaler J, Glöckner F-O, Schönhuber W, Amann R. Fluorescence in situ hybridization (FISH) with
652 rRNA-targeted oligonucleotide probes. In: *Methods in Microbiology*). Academic Press (2001).

653
654 56. Schindelin J*, et al.* Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9**,
655 676-682 (2012).

656
657 57. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data.
658 *Bioinformatics* **30**, 2114-2120 (2014).

659
660 58. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. Using SPAdes De Novo Assembler.
661 *Current Protocols in Bioinformatics* **70**, e102 (2020).

662
663 59. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome
664 assemblies. *Bioinformatics* **29**, 1072-1075 (2013).

665
666 60. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068-2069 (2014).

667
668 61. Galperin MY, Makarova KS, Wolf YI, Koonin EV. Expanded microbial genome coverage and
669 improved protein family annotation in the COG database. *Nucleic Acids Research* **43**, D261-D269
670 (2015).

671
672 62. Makarova KS, Wolf YI, Koonin EV. Archaeal Clusters of Orthologous Genes (arCOGs): An Update
673 and Application for Analysis of Shared Features between Thermococcales, Methanococcales, and
674 Methanobacteriales. *Life (Basel)* **5**, 818-840 (2015).

675
676 63. Aramaki T*, et al.* KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive
677 score threshold. *Bioinformatics* **36**, 2251-2252 (2020).

678
679 64. Bateman A*, et al.* The Pfam protein families database. *Nucleic Acids Res* **32**, D138-141 (2004).

680
681 65. Haft DH, Selengut JD, White O. The TIGRFAMs database of protein families. *Nucleic Acids Res* **31**,
682 371-373 (2003).

683
684 66. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. dbCAN: a web resource for automated carbohydrate-
685 active enzyme annotation. *Nucleic Acids Res* **40**, W445-451 (2012).

686

687 67. Saier MH, Jr., Tran CV, Barabote RD. TCDB: the Transporter Classification Database for membrane
688 transport protein analyses and information. *Nucleic Acids Res* **34**, D181-186 (2006).

689
690 68. Sondergaard D, Pedersen CN, Greening C. HydDB: A web tool for hydrogenase classification and
691 analysis. *Sci Rep* **6**, 34212 (2016).

692
693 69. Jones P*, et al.* InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**,
694 1236-1240 (2014).

695
696 70. Altschul SF*, et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search
697 programs. *Nucleic Acids Res* **25**, 3389-3402 (1997).

698
699 71. Buchfink B, Reuter K, Drost HG. Sensitive protein alignments at tree-of-life scale using DIAMOND.
700 *Nat Methods* **18**, 366-368 (2021).

701
702 72. Wörmer L, Lipp JS, Hinrichs K-U. Comprehensive analysis of microbial lipids in environmental
703 samples through HPLC-MS protocols. In: *Hydrocarbon and Lipid Microbiology Protocols:*
704 *Petroleum, Hydrocarbon and Lipid Analysis* (eds McGenity TJ, Timmis KN, Nogales B). Springer
705 Berlin Heidelberg (2015).

706
707 73. Sturt HF, Summons RE, Smith K, Elvert M, Hinrichs K-U. Intact polar membrane lipids in
708 prokaryotes and sediments deciphered by high-performance liquid chromatography/electrospray
709 ionization multistage mass spectrometry—new biomarkers for biogeochemistry and microbial
710 ecology. *Rapid Communications in Mass Spectrometry* **18**, 617-628 (2004).

711
712 74. Pluskal T, Castillo S, Villar-Briones A, Orešič M. MZmine 2: Modular framework for processing,
713 visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics*
714 **11**, 395 (2010).

715
716 75. Nothias L-F*, et al.* Feature-based molecular networking in the GNPS analysis environment. *Nature*
717 *Methods* **17**, 905-908 (2020).

718
719 76. Wang M*, et al.* Sharing and community curation of mass spectrometry data with Global Natural
720 Products Social Molecular Networking. *Nature Biotechnology* **34**, 828-837 (2016).

721
722 77. Watrous J*, et al.* Mass spectral molecular networking of living microbial colonies. *Proceedings of*
723 *the National Academy of Sciences* **109**, E1743 (2012).

724
725 78. Simone Z*, et al.* A Taxonomically-informed Mass Spectrometry Search Tool for Microbial
726 Metabolomics Data. *bioRxiv*, 2023.2007.2020.549584 (2023).

727

79. Smoot ME, Ono K, Ruscheinski J, Wang P-L, Ideker T. Cytoscape 2.8: New features for data integration and network visualization. *Bioinformatics* **27**, 431-432 (2010).

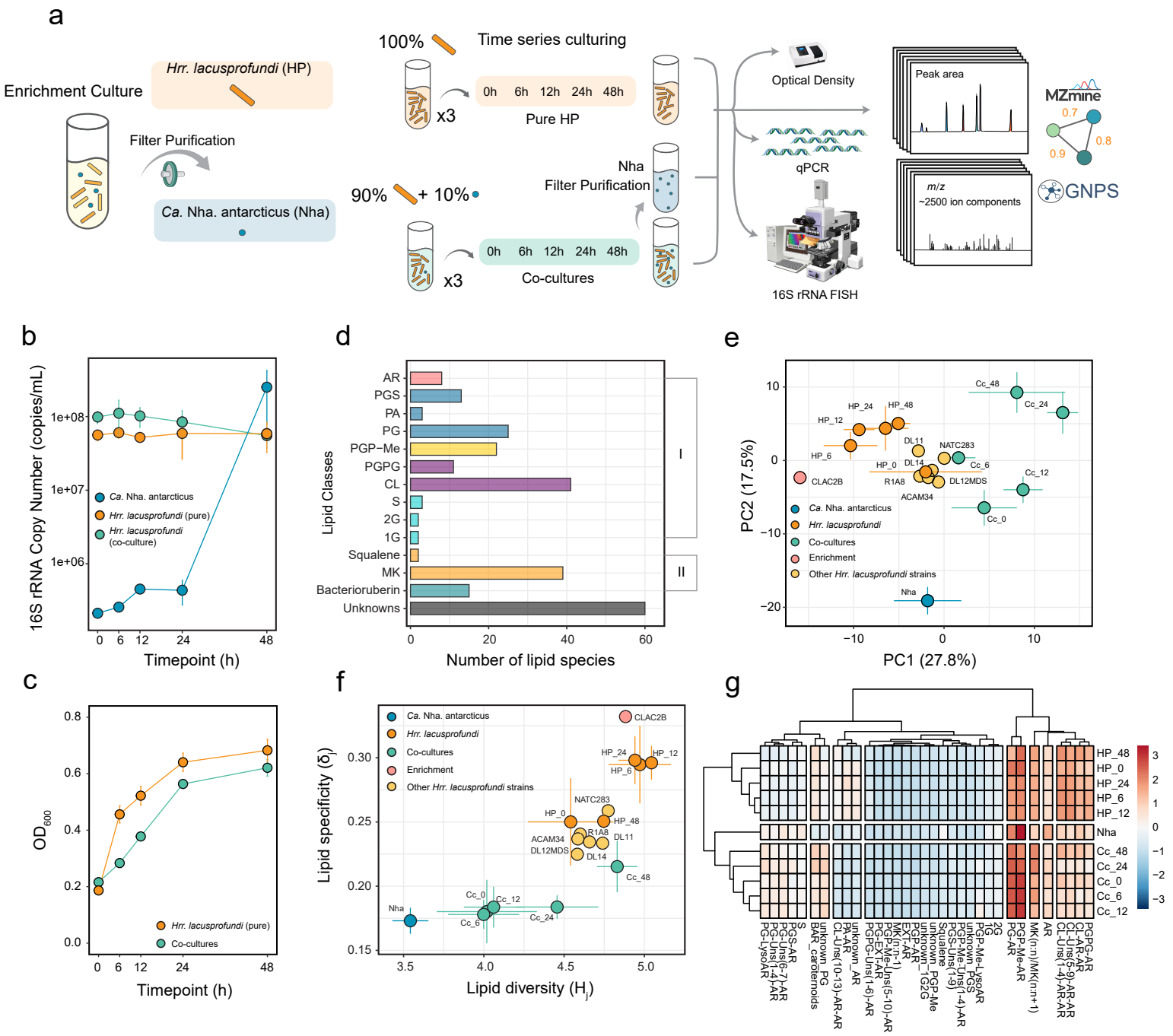
80. Shannon P*, et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research* **13**, 2498-2504 (2003).

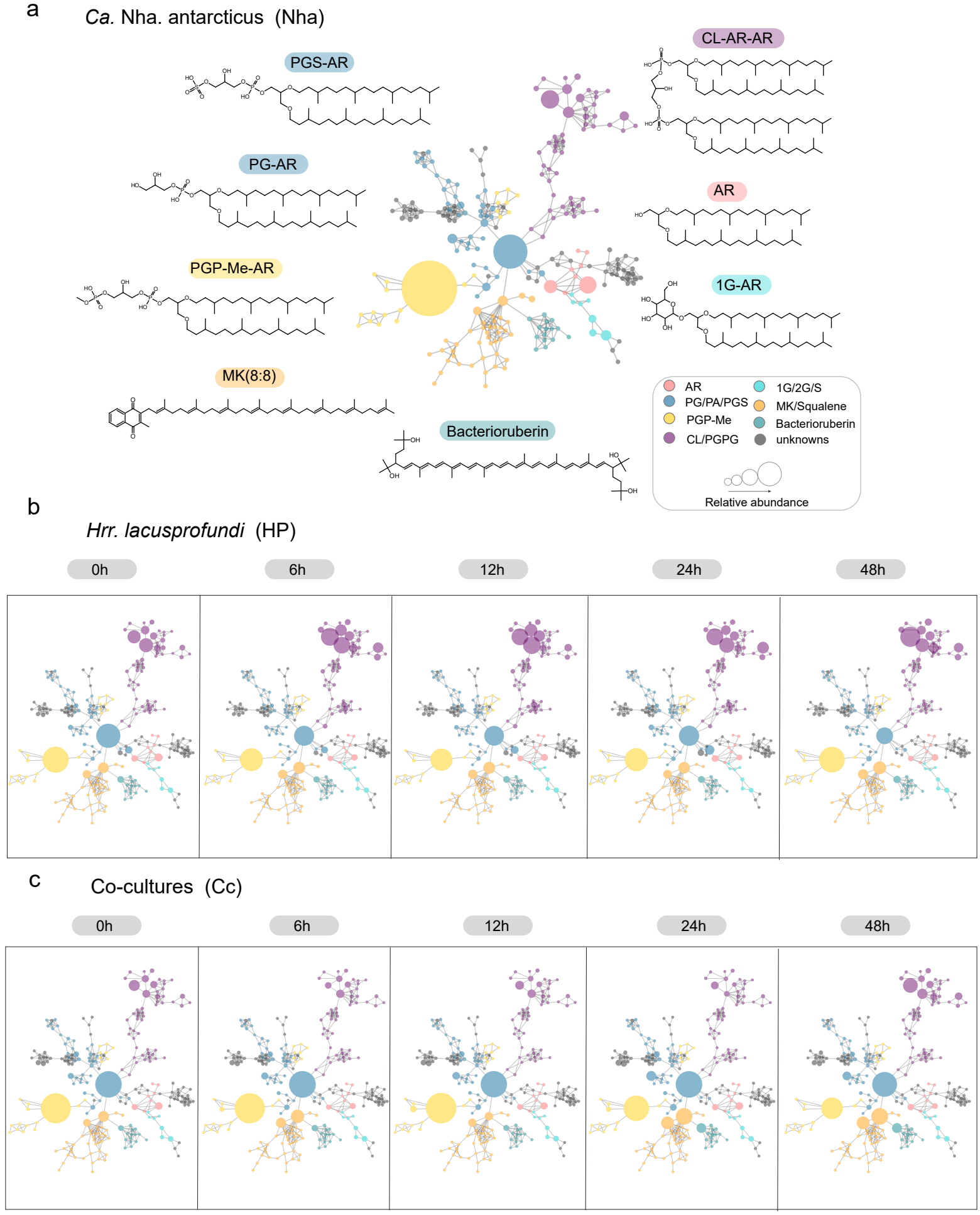
81. Ding S*, et al.* Changes in the membrane lipid composition of a Sulfurimonas species depend on the electron acceptor used for sulfur oxidation. *ISME Communications* **2**, 121 (2022).

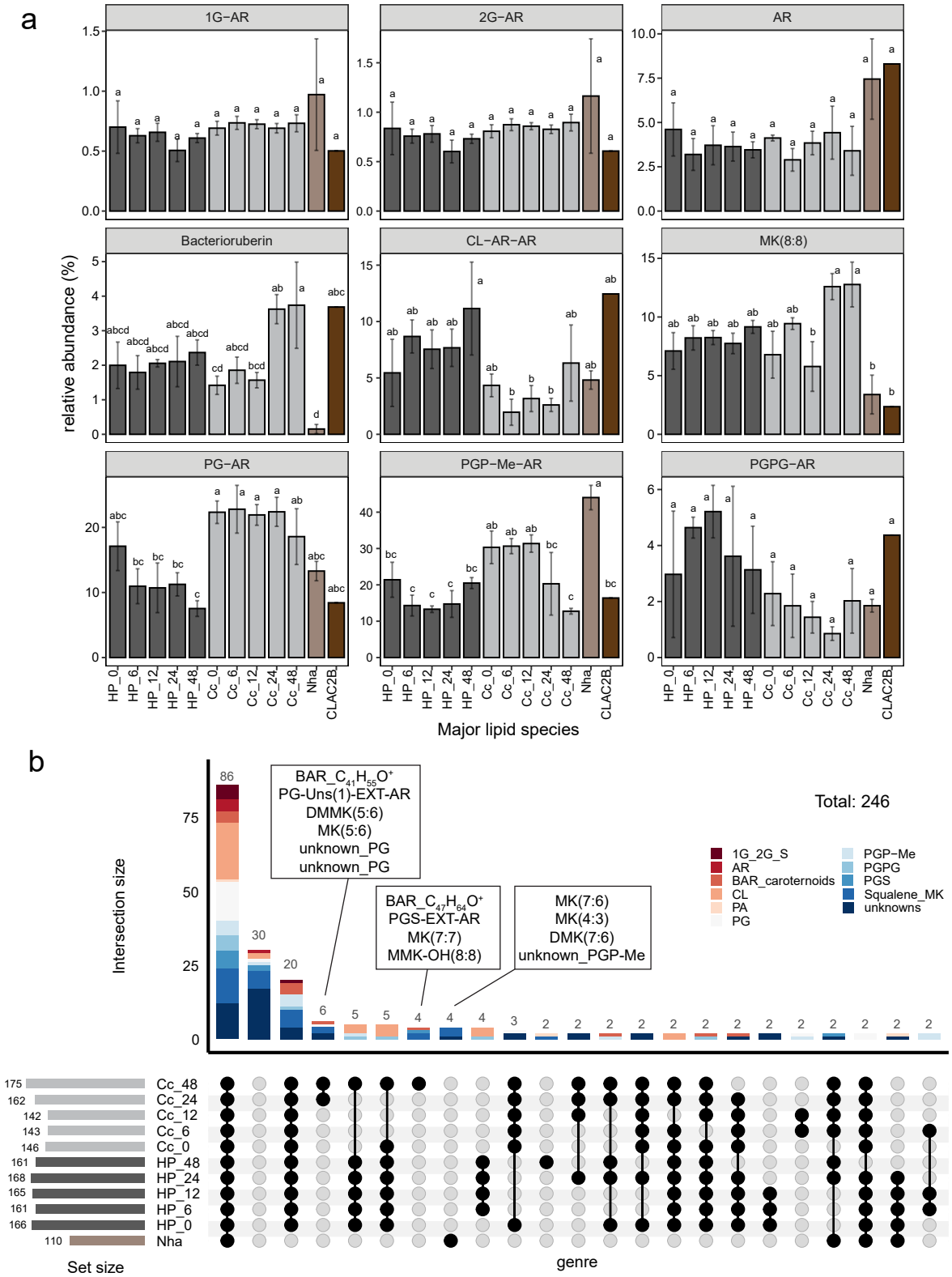
82. Li D, Heiling S, Baldwin IT, Gaquerel E. Illuminating a plant's tissue-specific metabolic diversity using computational metabolomics and information theory. *Proceedings of the National Academy of Sciences* **113**, E7610 (2016).

83. Lebret R, Collobert R. Word emdeddings through hellinger PCA. *arXiv preprint arXiv:13125542*, (2013).

84. Elling FJ*, et al.* Respiratory quinones in Archaea: phylogenetic distribution and application as biomarkers in the marine environment. *Environmental Microbiology* **18**, 692-707 (2016).


a
Enrichment Culture
Hrr. lacusprofundi (HP)
Filter Purification
Ca. Nha. antarcticus (Nha)
100%
Time series culturing
x3
0h 6h 12h 24h 48h
Pure HP
90% + 10%
Nha
Filter Purification
x3
0h 6h 12h 24h 48h
Co-cultures
Optical Density
qPCR
16S rRNA FISH
Peak area
m/z
~2500 ion components
MZmine
GNPS
b
16S rRNA Copy Number (copies/mL)
Timepoint (h)
Ca. Nha. antarcticus
Hrr. lacusprofundi (pure)
Hrr. lacusprofundi (co-culture)
c
OD600
Timepoint (h)
Hrr. lacusprofundi (pure)
Co-cultures
d
Lipid Classes
AR
PGS
PA
PG
PGP−Me
PGPG
CL
S
2G
1G
Squalene
MK
Bacterioruberin
Unknowns
I
II
Number of lipid species
e
PC2 (17.5%)
PC1 (27.8%)
Ca. Nha. antarcticus
Hrr. lacusprofundi
Co-cultures
Enrichment
Other Hrr. lacusprofundi strains
f
Lipid specificity (δj)
Lipid diversity (Hj)
Ca. Nha. antarcticus
Hrr. lacusprofundi
Co-cultures
Enrichment
Other Hrr. lacusprofundi strains
g
HP_48
HP_0
HP_24
HP_6
HP_12
Nha
Cc_48
Cc_24
Cc_0
Cc_6
Cc_12


a
Ca. Nha. antarcticus (Nha)
PGS-AR
PG-AR
PGP-Me-AR
MK(8:8)
Bacterioruberin
CL-AR-AR
AR
1G-AR
AR
PG/PA/PGS
PGP-Me
CL/PGPG
1G/2G/S
MK/Squalene
Bacterioruberin
unknowns
Relative abundance
b
Hrr. lacusprofundi (HP)
0h
6h
12h
24h
48h
c
Co-cultures (Cc)
0h
6h
12h
24h
48h


a
1G−AR
2G−AR
AR
Bacterioruberin
CL−AR−AR
MK(8:8)
PG−AR
PGP−Me−AR
PGPG−AR
relative abundance (%)
HP_0
HP_6
HP_12
HP_24
HP_48
Cc_0
Cc_6
Cc_12
Cc_24
Cc_48
Nha
CLAC2B
Major lipid species
b
Intersection size
Total: 246
BAR_C41H55O+
PG-Uns(1)-EXT-AR
DMMK(5:6)
MK(5:6)
unknown_PG
unknown_PG
BAR_C47H64O+
PGS-EXT-AR
MK(7:7)
MMK-OH(8:8)
MK(7:6)
MK(4:3)
DMK(7:6)
unknown_PGP-Me
1G_2G_S
AR
BAR_caroternoids
CL
PA
PG
PGP−Me
PGPG
PGS
Squalene_MK
unknowns
Set size
genre

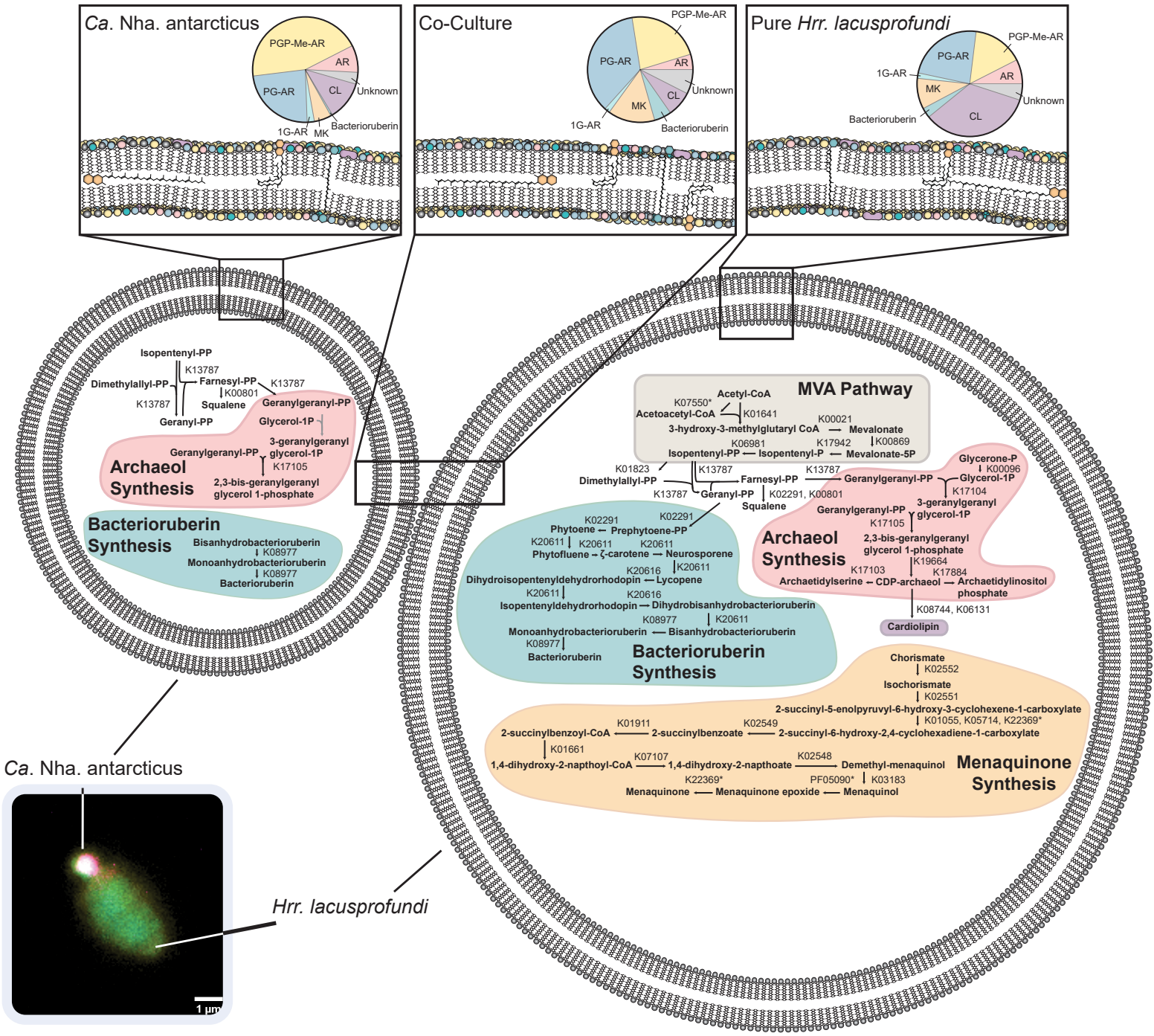

Ca. Nha. antarcticus
PGP-Me-AR
AR
PG-AR
CL
Unknown
1G-AR
MK
Bacterioruberin
Co-Culture
PGP-Me-AR
PG-AR
AR
MK
CL
Unknown
1G-AR
Bacterioruberin
Pure Hrr. lacusprofundi
PGP-Me-AR
PG-AR
1G-AR
AR
MK
Unknown
Bacterioruberin
CL
Isopentenyl-PP
K13787
Dimethylallyl-PP
Farnesyl-PP
K13787
K00801
K13787
Squalene
Geranyl-PP
Geranylgeranyl-PP
Glycerol-1P
Geranylgeranyl-PP
3-geranylgeranyl glycerol-1P
K17105
Archaeol Synthesis
2,3-bis-geranylgeranyl glycerol 1-phosphate
Bacterioruberin Synthesis
Bisanhydrobacterioruberin
K08977
Monoanhydrobacterioruberin
K08977
Bacterioruberin
MVA Pathway
Acetyl-CoA
K07550*
Acetoacetyl-CoA
K01641
K00021
3-hydroxy-3-methylglutaryl CoA
Mevalonate
K06981
K17942
K00869
Isopentenyl-PP
Isopentenyl-P
Mevalonate-5P
K01823
K13787
K13787
Dimethylallyl-PP
Farnesyl-PP
K13787
Geranylgeranyl-PP
Glycerone-P
K00096
Glycerol-1P
K13787
Geranyl-PP
K02291, K00801
Squalene
Geranylgeranyl-PP
K17104
3-geranylgeranyl glycerol-1P
K17105
Archaeol Synthesis
2,3-bis-geranylgeranyl glycerol 1-phosphate
K17103
K19664
K17884
Archaetidylserine
CDP-archaeol
Archaetidylinositol phosphate
K08744, K06131
Cardiolipin
K02291
K02291
Phytoene
Prephytoene-PP
K20611
K20611
K20611
Phytofluene
ζ-carotene
Neurosporene
K20616
K20611
Dihydroisopentenyldehydrorhodopin
Lycopene
K20611
K20616
Isopentenyldehydrorhodopin
Dihydrobisanhydrobacterioruberin
K08977
K20611
Monoanhydrobacterioruberin
Bisanhydrobacterioruberin
K08977
Bacterioruberin
Bacterioruberin Synthesis
Chorismate
K02552
Isochorismate
K02551
2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate
K01055, K05714, K22369*
K01911
K02549
2-succinylbenzoyl-CoA
2-succinylbenzoate
2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate
K01661
K07107
K02548
1,4-dihydroxy-2-napthoyl-CoA
1,4-dihydroxy-2-napthoate
Demethyl-menaquinol
K22369*
PF05090*
K03183
Menaquinone
Menaquinone epoxide
Menaquinol
Menaquinone Synthesis
Ca. Nha. antarcticus
Hrr. lacusprofundi
1 µm

Table 1. Detail of FISH Probes and PCR Primers used in this study

Name	**Purpose**	**Target Organism**	**Sequence**
Lacus-FISH-Cy3	FISH Probe	*Hrr. lacusprofundi*	/5Cy3/TTATTACAGTCGACGCTGG TGAGATGTCCG
Nha-FISH-Cy5	FISH Probe	*Ca.* Nha. antarcticus	/5Cy5/GTGTATCCCAGAGCATTCG
H_lac_qpcrF	qPCR Primer	*Hrr. lacusprofundi*	GGATTGTGCCAAAAGCTCCG
H_lac_qpcrR	qPCR Primer	*Hrr. lacusprofundi*	ACTCTCATGACCCGTACCGA
Nanohalo_qpcrF	qPCR Primer	*Ca.* Nha. antarcticus	ACTTAAAGGAATTGACGGGGG
Nanohalo_qpcrR	qPCR Primer	*Ca.* Nha. antarcticus	CATGCAGCTCCTCTCAGCG