INTRODUCTION

Bacteriohopanepolys (BHPs) are the prokaryotic surrogate of sterols and play an important role in the membranes of bacteria (Rohmer et al., 1979; Sáenz et al., 2012). BHPs and their degradation product, (geo)hopanoids, are common biomarkers in natural archives (Kimble et al., 1974; Ourisson & Albrecht, 1992; Quirk et al., 1984; Talbot & Farrimond, 2007). In addition to the common BHPs with a cyclized core skeleton, some bacteria produce 2-methylbacteriohopanepolys (2-MeBHPs; Figure 1), which are BHPs with an additional methyl group added to the A-ring at the second carbon position (Summons & Jahnke, 1990). Hopanoids

Received: 16 February 2021  |  Accepted: 10 July 2021
DOI: 10.1111/gbi.12465

ORIGINAL ARTICLE

The occurrence of 2-methylhopanoids in modern bacteria and the geological record

B. D. A. Naafs1  |  G. Bianchini2  |  F. M. Monteiro2  |  P. Sánchez-Baracaldo2

1Organic Geochemistry Unit, School of Chemistry and School of Earth Sciences, University of Bristol, Bristol, UK
2School of Geographical Sciences, University of Bristol, Bristol, UK

Correspondence
B. D. A. Naafs, Organic Geochemistry Unit, School of Chemistry and School of Earth Sciences, University of Bristol, BS8 1TS, UK.
Email: david.naafs@bristol.ac.uk

Abstract

The 2-methylhopanes (2-MeHops) are molecular fossils of 2-methylbacteriohopanepolys (2-MeBHPs) and among the oldest biomarkers on Earth. However, these biomarkers’ specific sources are currently unexplained, including whether they reflect an expansion of marine cyanobacteria. Here, we study the occurrence of 2-MeBHPs and the genes involved in their synthesis in modern bacteria and explore the occurrence of 2-MeHops in the geological record. We find that the gene responsible for 2-MeBHP synthesis (hpnP) is widespread in cyano- and α-proteobacteria, but absent or very limited in other classes/phyla of bacteria. This result is consistent with the dominance of 2-MeBHP in cyano- and α-proteobacterial cultures. The review of their geological occurrence indicates that 2-MeHops are found from the Paleoproterozoic onwards, although some Precambrian samples might be biased by drilling contamination. During the Phanerozoic, high 2-MeHops’ relative abundances (index >15%) are associated with climatic and biogeochemical perturbations such as the Permo/Triassic boundary and the Oceanic Anoxic Events. We analyzed the modern habitat of all hpnP-containing bacteria and find that the only one species coming from an undisputed open marine habitat is an α-proteobacterium acting upon the marine nitrogen cycle. Although organisms can change their habitat in response to environmental stress and evolutionary pressure, we speculate that the high sedimentary 2-MeHops’ occurrence observed during the Phanerozoic reflect α-proteobacteria expansion and marine N-cycle perturbations in response to climatic and environmental change.

KEYWORDS
2-methylhopanes, bacteria, biomarker, cyanobacteria, hopanoids

1 | INTRODUCTION

Bacteriohopanepolys (BHPs) are the prokaryotic surrogate of sterols and play an important role in the membranes of bacteria (Rohmer et al., 1979; Sáenz et al., 2012). BHPs and their degradation product, (geo)hopanoids, are common biomarkers in natural archives (Kimble et al., 1974; Ourisson & Albrecht, 1992; Quirk et al., 1984; Talbot & Farrimond, 2007). In addition to the common BHPs with a cyclized core skeleton, some bacteria produce 2-methylbacteriohopanepolys (2-MeBHPs; Figure 1), which are BHPs with an additional methyl group added to the A-ring at the second carbon position (Summons & Jahnke, 1990). Hopanoids
methylated at the C3 position (3-MeBHPs) also exist and might indicate methanotrophic bacteria (Farrimond et al., 2004; Zundel & Rohmer, 1985), but here we focus exclusively on the 2-MeBHPs/Hops. The 2-MeBHPs degrade over time and are preserved in the geological record as 2-methylhopanes (2-MeHops). First identified in oils (Seifert & Michael Moldowan, 1978), 2-MeHops are stable on geological time scales and among the oldest molecular fossils known on Earth (Summons et al., 1999). They are present in trace amounts or below detection limit in most natural samples through Earth’s history. However, elevated levels of 2-MeHops have been reported for some periods of Earth’s history (Kuypers et al., 2004; Summons et al., 1999).

The sedimentary occurrence of 2-MeHops was initially used to trace (aerobic) cyanobacteria (and photosynthesis and an oxidized atmosphere) based on initial culture data which indicated that a diverse range of aerobic cyanobacteria produce 2-MeBHPs (Summons et al., 1999). Kuypers et al., (2004) subsequently argued that the high relative abundance of 2-MeHops compared with regular hopanes (the 2-MeHops’ index (Summons et al., 1999)) found during the Oceanic Anoxic Events (OAEs) of the Cretaceous reflected an expansion of marine N₂-fixing cyanobacteria in response to the high degree of denitrification in anoxic oceans. Based on these landmark studies, 2-MeHops became a biomarker for marine (N₂-fixing) cyanobacteria and of widespread interest to the geobiology community (Schaefer et al., 2020; Sepúlveda et al., 2009; Xie et al., 2005).

However, this interpretation of 2-MeHops reflecting an expansion of aerobic marine (N₂-fixing) cyanobacteria has been challenged. In the last decade, advances in molecular biology showed that the anoxygenic α-proteobacteria (Rhodopseudomonas palustris) can also produce 2-MeBHPs (Rashby et al., 2007). In addition, the specific gene responsible for the methylation of BHPs at the C2 position (hpnP gene; Figure 1) was identified (Welander et al., 2010). Subsequent searches for the presence of the hpnP gene across bacterial phyla showed that hpnP is not exclusively found in cyanobacteria (Ricci et al., 2013; Welander et al., 2010). Recent studies have suggested that the high relative abundance of 2-MeHops during certain intervals of the sedimentary record can be used to identify periods when bacteria experienced stress (Garby et al., 2017b; Wu et al., 2015), potentially under anoxic ferrous conditions (Eickhoff et al., 2013), and/or that they indicate an environmental niche characterized by low oxygen and fixed nitrogen (Ricci et al., 2013).

**FIGURE 1** Synthesis pathway of 2-methylbacteriohopanetetrol and 2-methyldiplopterol, functionalized lipids that then degrade over time to 2-methylhopanes and preserved as such in the sedimentary record for hundreds of millions of years. Genes shown are squalene-hopene cyclase gene (Shc or hpnF) responsible for the cyclization of squalene and hpnP gene, responsible for the subsequent methylation at the C2 position.
Here we study 2-MeBHPs’ existence in modern bacteria, surveying a broad range of sequenced bacterial genomes, and review the 2-MeHops’ occurrence in the rock record to understand what drove the observed pattern of 2-MeHops’ occurrence during the Phanerozoic and Proterozoic.

2 | Methods

2.1 | Synthesis of 2-MeHops’ occurrence in cultures and the geological record

First, we compiled those studies that report the hopanoid distribution in cultures of bacteria (Table 1). This spans over 35 years of research, ranging from the classic Bisseret et al., (1985) study that first identified 2-MeHops in cultures of a cyanobacterium to the recent finding of a 2-MeHops producing acidobacterium (Sinninghe Damsté et al., 2017). We then compiled the occurrence of 2-MeHops in the geological record (Table 2). To the best of our knowledge, we report data from all peer-reviewed publications that report 2-MeHop data from well-dated sediments or oils. Data are presented as the maximum reported 2- MeHops’ index (Summons chromatogram (TIC), / m 191 and 205, or m/z 369 and 383) and hopanoids with different carbon chains (e.g., C30 vs. C31).

2.2 | Analysis of hpnF/hpnP genes in bacterial phyla

Distribution of 2-MeHops’ synthesis genes across bacterial genomes publicly available was assessed by BLAST searches. We downloaded the genome and proteome files of 14,624 completely sequenced bacterial strains from GenBank (Benson et al., 2017), through the NCBI genome FTP (Retrieved December 2017, from ftp://ftp.ncbi.nlm.nih.govGENOMES/). Accession numbers for all the genomes publicly available was assessed by BLAST searches. We downloaded the genome and proteome files of 14,624 completely sequenced bacterial strains from GenBank (Benson et al., 2017), through the NCBI genome FTP (Retrieved December 2017, from ftp://ftp.ncbi.nlm.nih.govGENOMES/). Accession numbers for all the genomes were available in the Supplementary Information (Table S1). We performed BLAST searches (Camacho et al., 2009) using as query sequences the hpnP gene (responsible for the methylation at the C2 position (Welander et al., 2010)) from Rhodopseudomonas palustris (accession number WP_011142311.1) and Gloecapsa violaceus (accession number WP_011142311.1) and blastp version 2.6.0+.

We identified hpnP ortholog genes (as previously described (Sánchez-Baracaldo et al., 2019) in a two-step process: First, we retained the first hit for each species, including hits with an e-value <10^-5. Then, we aligned these sequences using MAFFT v7.313 (Katoh & Standley, 2013) using the --maxiterate 1000 and --localpair options, and built a neighbor-joining tree in RapidNJ 2.3.2 (Simonsen et al., 2008). We identified “true” orthologs of hpnP using the tree topology.

We then built another alignment including the true orthologs of hpnP and removing positions with more than 85% gaps using AlignmentViewer online utility, http://sdsssdld.altervista.org/arklumps/AlignmentViewer/AlignmentViewer.html. Misaligned positions at the start and end of the alignment were also removed. Finally, this alignment was used to build a Bayesian gene tree (Figure 5) using MrBayes v3.2.7a (Ronquist & Huelsenbeck, 2003), under a mixed amino acid model prior, with a gamma rate heterogeneity model including invariant sites. Two independent runs were executed in parallel for 20,000,000 generations each, and convergence was assessed using the average standard deviation of split frequency statistic computed by MrBayes, as well as using TRACER v1.7.1 (Rambaut et al., 2018).

2.3 | Analysis of SSU rRNA from bacterial genomes and their habitat

To put into context the occurrence of hpnP in extant bacteria, we first estimated a bacterial tree (Figure 2) using SSU rRNA sequence data extracted from genomes using RNAmer v1.2 (Lagesen et al., 2007). We aligned SSU rRNA sequences using MAFFT v7.313 (options --ep 0.123 --nuc --adjustdirection=auto --maxiterate 1000 --retree 10). An approximate maximum-likelihood tree was built using FastTree v2.1.10 (Price et al., 2010) with options --gtr --nt --nosupport --spr 4 --gamma --fastest --no2nd. SSU % identity between every possible pair of strains was also computed using the alignment. We collapsed nodes in the tree based on the following 16S percentage identity thresholds: (i) Strains that possessed neither hpnF/shc nor hpnP and had not been tested for the production of 2-MeHop experimentally were collapsed using an 82.5% threshold, (ii) strains that possessed hpnF/shc, but not hpnP and had not been tested for the production of 2-MeHop experimentally were collapsed using a 90% threshold, and (iii) strains that possessed hpnP or had been tested for the production of 2-MeHop experimentally were collapsed using a 98.65% threshold. In Figure 2, the presence of a gene at a tip of the tree signifies that at least one of the strains underlie that tip possessed the gene. Habitat for the strains possessing the hpnP gene was obtained from the literature and/or the sequenced genome information (see supplementary information for taxa, habitat and literature source).

2.4 | Analysis of Synechococcus biomass for 2-MeHop

We analyzed biomass from Synechococcus elongatus PCC 7942 for its hopanoid content to verify specific culture data. The biomass was freeze-dried and extracted using a modified Bligh-Dyer protocol (Bligh & Dyer, 1959). The extraction solvent consisted of a combination of (i) 100 ml of aqueous phosphate buffer (pH 7.2), (ii) 250 ml of methanol (MeOH), and (iii) 125 ml dichloromethane (DCM). 15 ml of this solvent was used to extract 0.1 g of Synechococcus biomass. The biomass-solvent mixture was ultrasonicated for 10 min and centrifuged at 3000 rpm (5 min) after which the supernatant was collected. This was repeated four times and the supernatants combined. 16 ml of DCM and 16 ml of aqueous phosphate buffer were
| Phylum/Class   | Species Name                  | Strain                  | 2-MeBHP? | Reference | hpnP gene? |
|---------------|------------------------------|-------------------------|----------|-----------|------------|
| Cyanobacteria | Anacystis Montana            | CCAP 1405/3             | Yes      | 1,2       | Yes        |
| Cyanobacteria | Cyanothece                   | RCB4                    | Yes      | 1,3       | G.N.A      |
| Cyanobacteria | Gloeobacter sp.              |                         | Yes      | 4         | G.N.A      |
| Cyanobacteria | Gloeobacter violaceus        | ATCC 29082              | Yes      | 4         | Yes        |
| Cyanobacteria | Halothece sp.                | PCC 7418                | Yes*     | 4         | No         |
| Cyanobacteria | Synechococcus                | PCC 6301                | Yes*     | 4         | No         |
| Cyanobacteria | Synechococcus lividus        | ATCC 27180              | Yes      | 3,4       | Yes        |
| Cyanobacteria | Synechococcus sp.            | L 1402-1                | No       | 5         | No         |
| Cyanobacteria | Synechococcus sp.            | PCC 6907                | Yes      | 6         | G.N.A      |
| Cyanobacteria | Synechococcus elongatus      | PCC 7942                | No       | This study| No         |
| Cyanobacteria | Synechocystis sp.            | ATCC 27170              | No       | 5         | G.N.A      |
| Cyanobacteria | Synechocystis sp.            | ATCC 27178              | No       | 5         | No         |
| Cyanobacteria | Synechocystis sp.            | PCC 6714                | No       | 7         | No         |
| Cyanobacteria | Anabaena cylindrica          | ATCC 27899              | No       | 4         | No         |
| Cyanobacteria | Anabaena variabilis          |                         | No       | 5         | No         |
| Cyanobacteria | Calothrix anomala            |                         | Yes      | 3,4       | G.N.A      |
| Cyanobacteria | Calothrix sp.                | ATCC 27914              | No       | 5         | No         |
| Cyanobacteria | Calothrix sp.                | TSB                     | No       | 4         | G.N.A      |
| Cyanobacteria | Calothrix sp.                | NP                      | Yes      | 1         | G.N.A      |
| Cyanobacteria | Nostoc muscorum              | B 1452-12b              | Yes      | 8         | G.N.A      |
| Cyanobacteria | Nostoc muscorum              | PCC 6720                | Yes      | 9         | G.N.A      |
| Cyanobacteria | Nostoc punctiforme           | ATCC 291335             | Yes      | 10,11     | Yes        |
| Cyanobacteria | Nostoc punctiforme           | PCC 73102               | Yes      | 1,12      | Yes        |
| Cyanobacteria | Nostoc muscorum              | CCAP 1453/12            | Yes      | 1         | G.N.A      |
| Cyanobacteria | Scytonema hofmanni           | ATCC 29171              | Yes      | 1,5       | Yes        |
| Cyanobacteria | Lyngbya–Phormidium–Plectonema| ATCC 27894              | No       | 5         | Yes        |
| Cyanobacteria | Lyngbya–Phormidium–Plectonema| ATCC 27902              | No       | 5         | G.N.A      |
| Cyanobacteria | Oscillatoria amphigranulata  | OSU                     | Yes      | 1,3       | G.N.A      |
| Cyanobacteria | Oscillatoria limnetica       | (Y. Cohen)              | No       | 4         | G.N.A      |
| Cyanobacteria | Oscillatoria rubescens       | Traces                  | No       | 4         | No         |
| Cyanobacteria | Oscillatoria sp.             | ATCC 27935              | No       | 5         | G.N.A      |
| Cyanobacteria | Oscillatoria sp.             | GN                      | No       | 4         | G.N.A      |
| Cyanobacteria | Oscillatoria sp.             | (L. Stal)               | No       | 4         | G.N.A      |
| Cyanobacteria | Oscillatoria sp.             | (J. Bauld)              | No       | 4         | G.N.A      |
| Cyanobacteria | Phormidium                   | FPG4                    | No       | 3         | G.N.A      |
| Cyanobacteria | Phormidium                   | FPOS4                   | Yes      | 3         | G.N.A      |
| Cyanobacteria | Phormidium                   | RCG3                    | No       | 3,4       | G.N.A      |
| Cyanobacteria | Phormidium                   | RCG4                    | No       | 3,4       | G.N.A      |
| Cyanobacteria | Phormidium luridum           | UTEX 426 (CCAP 1462/2)  | Yes      | 1,3,4     | Yes        |
| Cyanobacteria | Phormidium sp.               | GN                      | No       | 4         | G.N.A      |
| Cyanobacteria | Phormidium sp.               | OSS3                    | No       | 4         | G.N.A      |
| Cyanobacteria | Phormidium sp.               | OSS4                    | Yes      | 1,3,4     | G.N.A      |
| Cyanobacteria | Phormidium sp.               | RCG                      | Yes      | 1         | G.N.A      |

(Continues)
| Phylum/Class | Species Name       | Strain     | 2-MeBHP? | Reference | hpnpP gene? |
|-------------|--------------------|------------|----------|-----------|-------------|
| Cyanobacteria | Phormidium sp.    | RCO        | Yes      | 1         | G.N.A       |
| Cyanobacteria | Schizothrix sp.   |            | No       | 4         | G.N.A       |
| Cyanobacteria | Spirulina sp.     |            | No       | 5         | G.N.A       |
| Cyanobacteria | Spirulina sp.     | IFP        | No       | 1         | G.N.A       |
| Cyanobacteria | Prochlorothrix hollandica | CCAP 1490/1 | Yes | 1,13 | Yes |
| Cyanobacteria | Chlorogloeopsis sp. | LA         | No       | 1         | G.N.A       |
| Cyanobacteria | Chlorogloeopsis fritschii | ATCC 27193 | Yes | 1,3,4 | Yes |
| Cyanobacteria | Fischerella sp.   | ATCC 29538 | No       | 3,4       | No          |
| Cyanobacteria | Fischerella sp.   |            | No       | 5         | G.N.A       |
| Cyanobacteria | Crocosphaera sp.  | WH8501     | No       | 1         | No          |
| Cyanobacteria | Synechococcus     | WH8102     | No       | 1         | No          |
| Cyanobacteria | Trichodesmium sp. | IMS-101    | No       | 1         | No          |
| Cyanobacteria | Prochlorococcus marinus | MIT 9312   | No       | 1         | No          |
| Cyanobacteria | Prochlorococcus marinus | MIT 9313   | No       | 1         | No          |
| Cyanobacteria | Prochlorococcus marinus | CCMP 1378 / MED4 | No | 1 | No |
| Cyanobacteria | Microcystis sp.   | CCAP 1450/13 | No | 1 | G.N.A |
| Cyanobacteria | Microcystis sp.   | 110        | No       | 1         | G.N.A       |
| Cyanobacteria | Synechocystis sp. | CCAP 1480/4 | No | 1 | G.N.A |
| Cyanobacteria | Synechocystis sp. | PCC 6803   | No       | 1         | No          |
| Cyanobacteria | Anabena cylindrica | ATTC culture | No | 1 | G.N.A |
| Cyanobacteria | Chroococcidiopsis sp. |           | Yes     | 1         | G.N.A       |
| Cyanobacteria | Microcystis aeruginosa | PCC 7808 | No | 1 | G.N.A |
| α-Proteobacteria | Rhodopseudomonas palustris | TIE-1 | Yes | 14–16 | Yes |
| α-Proteobacteria | Rhodopseudomonas palustris | CGA009 | Yes | 14 | Yes |
| α-Proteobacteria | Rhodopseudomonas palustris | BisB18 | Yes | 12 | Yes |
| α-Proteobacteria | Rhodopseudomonas palustris | BisA53 | Yes | 12 | Yes |
| α-Proteobacteria | Rhodopseudomonas palustris | BisB5 | Yes | 12 | Yes |
| α-Proteobacteria | Rhodopseudomonas palustris | HaA2 | Yes | 12 | Yes |
| α-Proteobacteria | Rhodopseudomonas acidiphila | 7050 | No | 5 | No |
| α-Proteobacteria | Rhodopseudomonas acidiphila | 10050 | No | 5 | G.N.A |
| α-Proteobacteria | Bradyrhizobium sp. | BTAi1 | Yes | 17 | Yes |
| α-Proteobacteria | Bradyrhizobium japonicum | USDA110 | Yes | 17 | Yes |
| α-Proteobacteria | Bradyrhizobium diazoefficiens | BTAi1 | Yes | 18 | Yes |
| α-Proteobacteria | Beijerinckia indica | ATCC 9039 | Yes | 19 | Yes |
| α-Proteobacteria | Beijerinckia mobilis | DSM 2326 | Yes | 19 | Yes |
| α-Proteobacteria | Methylobacterium organophilum | NCIB 11278 | Yes | 8 | G.N.A |
| α-Proteobacteria | Azotobacter vinelandii | DSM 2289 | No | 19 | G.N.A |
| α-Proteobacteria | Zymomonas mobilis | ATCC 29191 & ATCC 31821 | No | 20 | No |
| α-Proteobacteria | Rhodococcus vannielii |            | No | 5 | No |
| α-Proteobacteria | Rhodospirillum rubrum |            | No | 5 | No |
| α-Proteobacteria | Rhodospirillum rubrum | Ha | No | 5 | G.N.A |
| α-Proteobacteria | Gluconobacter oxydans |            | No | 5 | No |
| α-Proteobacteria | Acetobacter pasteurianus |            | No | 21 | No |

(Continues)
| Phylum/Class | Species Name                      | Strain | 2-MeBHP? | Reference | hpnP gene? |
|-------------|----------------------------------|--------|----------|-----------|------------|
| α-Proteobacteria | Hyphomicrobium sp.          | X      | No       | 5         | G.N.A      |
| α-Proteobacteria | Azotobacter chroococcum CIPI   | CIPI   | No       | 5         | G.N.A      |
| α-Proteobacteria | A. aceti subsp. aceti         | NCIB 8621 | No     | 5         | No         |
| α-Proteobacteria | A. aceti subsp. liquefaciens   | NCIB 9418 | No     | 5         | G.N.A      |
| α-Proteobacteria | A. aceti subsp. xylinum        | NCIB 41 12 | No     | 5         | G.N.A      |
| α-Proteobacteria | A. aceti subsp. xylinum        | R 2277  | No       | 5         | G.N.A      |
| α-Proteobacteria | A. pasteurianus subsp. estunensis | NCIB 893 | No     | 5         | G.N.A      |
| α-Proteobacteria | A. pasteurianus subsp. lovaniensis | NCIB 8620 | No     | 5         | G.N.A      |
| α-Proteobacteria | A. pasteurianus subsp. orleanensis | NCIB 6426 | No     | 5         | G.N.A      |
| α-Proteobacteria | A. pasteurianus subsp. pasteurianus | NCIB 429 & | No     | 5         | G.N.A      |
| α-Proteobacteria | A. peroxydans                   | NCIB 8087 | No     | 5         | G.N.A      |
| α-Proteobacteria | A. peroxydans                   | NCIB 8618 | No     | 5         | No         |
| α-Proteobacteria | Rhodoplanes oryzae             | JA793   | Yes      | 22        | G.N.A      |
| α-Proteobacteria | Rhodoplanes elegans            | DSM11907 | Yes     | 22        | Yes        |
| α-Proteobacteria | Rhodoplanes piscinae           | JA266   | Yes      | 22        | Yes        |
| α-Proteobacteria | Rhodoplanes pokkaliisoli       | JA415   | Yes      | 22        | G.N.A      |
| α-Proteobacteria | Rhodoplanes roseus             | DSM5909 | Yes     | 22        | Yes        |
| α-Proteobacteria | Methylorubrum extorquens       |        | Yes      | 23        | Yes        |
| α-Proteobacteria | Methyllobacterium fujisawaense  |        | Yes      | 24        | G.N.A      |
| α-Proteobacteria | Methyllobacterium mesophilicum  | ATCC 29983 | Yes  | 24        | Yes        |
| α-Proteobacteria | Nitrobacter vulgaris           | AB1     | Yes      | 25        | Yes        |
| γ-Proteobacteria | Methylcystis parvus            | OBBP    | No       | 5         | Yes        |
| γ-Proteobacteria | Methyllosinus sporum           | 5       | No       | 5         | No         |
| γ-Proteobacteria | Methyllosinus trichosporium    | PG      | No       | 5         | G.N.A      |
| γ-Proteobacteria | Methyllosinus trichosporium    | OB3b    | No       | 5         | No         |
| β-Proteobacteria | Nitrosomonas europaea          |         | No       | 5         | No         |
| β-Proteobacteria | B. caryophylli                 | DSM 50341 | No    | 26        | No         |
| β-Proteobacteria | B. gladioli                    | DSM 4285 | No       | 26        | No         |
| β-Proteobacteria | B. cepacia                     | G-21019 | No       | 26        | G.N.A      |
| β-Proteobacteria | B. pseudomallei                | G-22009 | No       | 26        | G.N.A      |
| β-Proteobacteria | Burkholderia cenocepa          |         | No       | 27,28     | No         |
| β-Proteobacteria | Burkholderia pseudomallei      |         | No       | 26        | No         |
| β-Proteobacteria | Pseudomonas cepacia            | Berkeley 382 | No     | 5         | G.N.A      |
| δ-Proteobacteria | Geobacter metallireducens      | GS-15 (DSM 7210) | No   | 29        | No         |
| δ-Proteobacteria | Geobacter sulfurreducens       | PCA (DSM 12127; ATCC 51573) | No | 29       | No         |
| γ-Proteobacteria | Methylobacter capsulatus       |         | No       | 5         | No         |
| γ-Proteobacteria | Methylobacter capsulatus       | TRMC    | No       | 5         | No         |
| γ-Proteobacteria | Azotobacter vinelandii         | CCM 289 | No       | 5         | G.N.A      |
| γ-Proteobacteria | P. amygdali / P. chlororaphis  | G-21032 | No       | 26        | G.N.A      |
| Actinobacteria   | Streptomyces chartreus         | NRLL 3882 | No       | 5         | No         |

(Continues)
| Phylum/Class | Species Name                        | Strain        | 2-MeBHP? | Reference | hpnP gene? |
|-------------|------------------------------------|---------------|---------|-----------|------------|
| Actinobacteria | Streptomyces sp.                     | G 1815        | No      | 5         | G.N.A      |
| Actinobacteria | Streptomyces sp.                     | 4_6609        | No      | 5         | G.N.A      |
| Actinobacteria | Streptomyces coelicolor              | A3(2)         | No      | 30        | No         |
| Actinobacteria | Frankia alni                         | HFPAr13       | No      | 31        | G.N.A      |
| Acidobacteria | Edaphobacter aggregans Wbg-1         | DSM 19364     | No      | 32        | No         |
| Acidobacteria | Edaphobacter modestus Jbg-1          | DSM 18101     | No      | 32        | No         |
| Acidobacteria | Acidobacterium capsulatum 161        | DSM 11244     | No      | 32        | No         |
| Acidobacteria | Ococcusiobacter savannea A2-1c       | DSM 25170     | No      | 32        | No         |
| Acidobacteria | Ococcusiobacter riparius 277         | DSM 25168     | No      | 32        | No         |
| Acidobacteria | Ococcusiobacter riparius 307         | DSM 25169     | No      | 32        | No         |
| Acidobacteria | Acidobacteriaeae bacterium A2-4c     |              | No      | 32        | No         |
| Acidobacteria | Acidicapsa borealis KA1              | DSM 23886     | No      | 32        | No         |
| Acidobacteria | Acidicapsa ligni WH120               | LMG 26244     | No      | 32        | No         |
| Acidobacteria | Acidicapsa sp. CEI                   |              | No      | 32        | No         |
| Acidobacteria | Ca. Koribacter versatilis Ellin345   | DSM 22529     | Yes     | 32        | Yes        |
| Acidobacteria | Terriglobus roseus KBS 63            | DSM 18391     | No      | 32        | No         |
| Acidobacteria | Granulicella pectinivorans TPB6011   | DSM 21001     | No      | 32        | No         |
| Acidobacteria | Granulicella aggregans TPB6028       | DSM 25274     | No      | 32        | No         |
| Acidobacteria | Granulicella paludicola LCBR         |              | No      | 32        | No         |
| Acidobacteria | Granulicella paludicola OB1010       | DSM 22464     | No      | 32        | No         |
| Acidobacteria | Broycella elongata SN10              | DSM 22489     | No      | 32        | No         |
| Acidobacteria | Telmatobacter bradus TPB6017         | DSM 23630     | No      | 32        | No         |
| Acidobacteria | Telmatobacter sp. 15–8A             |              | No      | 32        | No         |
| Acidobacteria | Paludibaculum fermentans P105        | DSM 26340     | No      | 32        | No         |
| Acidobacteria | Ca. Solibacter usitatus Ellin6076    |              | No      | 32        | No         |
| Acidobacteria | Bryobacter aggregatus MPL3           | DSM 18758     | No      | 32        | No         |
| Acidobacteria | Bryobacter aggregatus MPL1011        |              | No      | 32        | No         |
| Acidobacteria | Chloracidobacterium thermophilum B   | ATCC BAA-2647 | No      | 32        | No         |
| Acidobacteria | Pyrinomonas methylalipathogenes K22  | DSM 25857     | No      | 32        | No         |
| Acidobacteria | Blastocatella fastidiosa A2-16       | DSM 25172     | No      | 32        | No         |
| Acidobacteria | Brevitalea aridisoli Ac_11_E3        | DSM 27934     | No      | 32        | No         |
| Acidobacteria | Stenotrophobacter terrae Ac_28_D10   | DSM 26560     | No      | 32        | No         |
| Acidobacteria | Aribacter kavangonensis Ac_23_E3     | DSM 26558     | No      | 32        | No         |
| Acidobacteria | Aribacter famidurans A22_HD_4H       | DSM 26555     | No      | 32        | No         |
| Acidobacteria | Vicinamibacter silvestris Ac_5_C6    | DSM 29464     | No      | 32        | No         |
| Acidobacteria | Luteitea pratisensis HEG_ -6_39      | DSM 100886    | No      | 32        | No         |
| Acidobacteria | Holophaga foetida TMBS4              | DSM 6591      | No      | 32        | No         |
| Acidobacteria | Geothrix fermentans H-5              | DSM 14018     | No      | 32        | No         |
| Acidobacteria | Thermotomaculum hydrothermal AC55    | DSM 24660     | No      | 32        | No         |
Our genomic analyses show that the \textit{hpnF} gene is relatively widespread among a range of bacterial phyla, while the \textit{hpnP} gene is less common (Figure 2). \textit{HpnP} is predominantly found in cyanobacteria and \textit{α}-proteobacteria (mostly Rhizobiales). One acidobacterium (\textit{Candidatus Koribacter versatilis}) and three closely related actinobacteria (\textit{Asanoa ferruginea}, \textit{Streptomyces purpurigenescleroticus}, and \textit{Streptomyces sp. CB02959}) possess the \textit{hpnP} gene (as well as \textit{hpnF} gene). The \textit{hpnP} gene is absent in all sequenced firmicutes, \textit{β}-, \textit{γ}-, and \textit{δ}-proteobacteria.

The literature survey of bacterial cultures indicates that with one exception (an acidobacterium), all cultures that reported 2-MeBHPs belong to the phyla of cyanobacteria and \textit{α}-proteobacteria (Figure 2). Our HT-GC-MS analysis of the \textit{Synechococcus elongatus} PCC 7942 biomass found no distinct peaks in the m/z 191 and 205 partial mass spectra, the masses characteristic for hopanoids and 2-methylhopanoids respectively (Figure 3).

### 3.2 2-MeHops in the rock record

Our synthesis on 2-MeHops’ occurrence in the geological record (Figure 4) indicates that these lipids have been present since at least the paleo-Proterozoic. The 2-MeHops’ distributions in ancient samples from the Mt McRae Shale (2.5 Gyr) and Wollongorang FM (1.7 Gyr) (Summons et al., 1999), Matinenda Fm (2.45 Gyr) (Welander et al., 2010), Transvaal Supergroup (2.5 Gyr) (Waldbauer et al., 2009), and Archaean rocks from the Pilbara Craton (~2.7 Gyr) (e.g. Brocks et al., 1999; Eigenbrode et al., 2008) are likely biased due to contamination (French et al., 2015) and not shown here. The occurrence of 2-MeHops in some of the Meso- and Paleoproterozoic samples might also have been contaminated, but at this point, we cannot objectively assess this so report the data as originally published.

In total, we compiled data from around 80 studies that reported 2-MeHops’ occurrence, scattered throughout the last 1800 Myr with...
### TABLE 2 Compilation of the maximum relative abundance of 2-MeHop in sediments and oils through Earth’s history

| Age (Myr) | Max. 2-MeHop index (%) | Location | Event | Environment | Reference |
|-----------|------------------------|----------|-------|-------------|-----------|
| 1         | 0.2                    | Site 1084|       | Marine      | 1         |
| 5.2       | 0.4                    | Gessosso Sofefera Fm. |       | Marine      | 1         |
| 7         | 0.9                    | Monterey Fm. | Oils/Bitumen | Marine | 2         |
| 14        | 1.5                    | Monterey Fm. |       | Marine      | 1         |
| 20.5      | 5                      | Monterey Fm. | Oils/Bitumen | Marine | 2         |
| 33.5      | 0.4                    | Menilite Fm. |       | Marine      | 1         |
| 49        | 1.43                   | Green River Fm. | PETM (likely Cret. origin) | Lacustrine | 3         |
| 56        | 75.5                   | South Dover Bridge |       | Marine      | 4         |
| 63        | 0.3                    | Taranaki basin | Oils/Bitumen | Marine | 2         |
| 66        | 5                      | Site M0077A | K/Pg boundary | Marine | 5         |
| 66.5      | 9.8                    | Fish Clay | K/Pg boundary | Marine | 6         |
| 85        | 2                      | Site 959 |       | Marine      | 1         |
| 93.7      | 14                     | Covelo | OAE 2 | Marine | 7         |
| 93.7      | 15                     | Monte Nasseba | OAE 2 | Marine | 7         |
| 93.7      | 36                     | Oued Bahloul | OAE 2 | Marine | 7         |
| 93.7      | 40                     | Site 144 | OAE 2 | Marine | 8         |
| 93.7      | 45                     | Site 367 | OAE 2 | Marine | 8         |
| 93.7      | 33                     | Site 603 | OAE 2 | Marine | 8         |
| 93.7      | 7.5                    | Bahloul Fm. | OAE 2 | Marine | 9         |
| 100       | 0.9                    | Toolebuc Fm. |       | Marine | 1         |
| 102.5     | 0.6                    | Julia Ck Fm. | Oils/Bitumen | Marine | 2         |
| 120       | 43                     | Site 463 | OAE 1a | Marine | 10        |
| 120       | 21                     | Cismon Core | OAE 1a | Marine | 10        |
| 120       | 33                     | Site 1207 | OAE 1a | Marine | 8         |
| 120       | 68                     | Site 1213 | OAE 1a | Marine | 8         |
| 120       | 3.1                    | Livello Selli | OAE 1a | Marine | 9         |
| 137       | 2                      | Vergol-Morenas section |       | Marine | 11        |
| 150       | 18.8                   | Calcaires en Plaquettes Fm. | Mircobial mat | Marine | 1         |
| 155       | 4.5                    | Vulcan sub-basin oil | Oils/Bitumen | Marine | 2         |
| 155       | 4.5                    | Barrow sub-basin oil | Oils/Bitumen | Marine | 2         |
| 155       | 2.5                    | North Sea oil | Oils/Bitumen | Marine | 2         |
| 155       | 2.7                    | Kimmeridge Clay |       | Marine | 9         |
| 156       | 0.8                    | Schistes Cartons Fm. |       | Marine | 1         |
| 183       | 5                      | Val Varea | Toarcian OAE | Marine | 12        |
| 183       | 14                     | Rizapol | Toarcian OAE | Marine | 12        |
| 183       | 22                     | Longarone | Toarcian OAE | Marine | 12        |
| 183       | 16                     | Monte Mangart | Toarcian OAE | Marine | 12        |
| 183       | 2.6                    | Jet Rock Fm. | Toarcian OAE | Marine | 13        |
| 183       | 1.8                    | UK | Toarcian OAE | Marine | 9         |
| 183       | 2.4                    | Italy | Toarcian OAE | Marine | 9         |
| 190       | 10.49                  | Towaco Fm. |       | Lacustrine | 14        |
| 201.4     | 28.68                  | Peril and Sandlands Fm. | End-Triassic | Marine | 15        |
| 239       | 41                     | Serpiano oil Shale | Black shale | Marine | 9,16      |
| 248       | 43                     | Feixianguan Fm. | Perm/Tri bound. | Marine | 17        |
| 249       | 40                     | Chaohu section | Perm/Tri bound. | Marine | 18        |
| Age (Myr) | Max. 2-MeHOP index (%) | Location | Event | Environment | Reference |
|----------|------------------------|----------|-------|-------------|-----------|
| 251      | 47                     | Qingyan section | Perm/Tri bound. | Marine | 18 |
| 252      | 32.54                  | Meishan sections | Perm/Tri bound. | Marine | 19,20 |
| 253      | 11                     | Kap Stosch section | Perm/Tri bound. | Marine | 21 |
| 256      | 4.6                    | Marl Slate |                | Marine | 9 |
| 285      | 0.5                    | Phosphoria Fm. oil | Oils/Bitumen | Marine | 2 |
| 337.5    | 9.4                    | Milligans Fm. | Oils/Bitumen | Marine | 2 |
| 362      | 6.5                    | Pando X-1 Core |                | Marine | 22 |
| 368      | 5.5                    | Irish Gulf (Appalachian Basin) |                | Marine | 22 |
| 448      | 11.3                   | Anticosti Island |                | Marine | 23 |
| 455      | 21                     | Cominco SS-9 core |                | Marine | 24 |
| 463      | 11.7                   | Goldwyer Fm. |                | Marine | 25 |
| 472.5    | 5.3                    | Horn Valley Fm. | Oils/Bitumen | Marine | 2 |
| 480      | 12                     | Kirtonryggen Fm. |                | Marine | 26 |
| 507.5    | 2.4                    | Inca Fm. | Oils/Bitumen | Marine | 2 |
| 535      | 32.4                   | Observatory Hills Fm. | Oils/Bitumen | Marine | 2 |
| 535      | 9.9                    | Ouldburra Fm. | Oils/Bitumen | Marine | 2 |
| 535      | 9.6                    | Ara5-6 oils | Oils/Bitumen | Marine | 27 |
| 541      | 10.2                   | Ara4 oils | Oils/Bitumen | Marine | 27 |
| 543      | 7.8                    | Ara1-Ara3 oils | Oils/Bitumen | Marine | 27 |
| 550      | 11.2                   | Athel Fm. | Oils/Bitumen | Marine | 2 |
| 550      | 10.47                  | Nepa-Botuoba oils | Oils/Bitumen | Marine | 28 |
| 560      | 6.7                    | Ungoolya Fm. | Oils/Bitumen | Marine | 2 |
| 560      | 4.9                    | Pertatataka Fm. | Oils/Bitumen | Marine | 2 |
| 570      | 15.9                   | Shuram rocks |                | Marine | 29 |
| 600      | 6                      | Khufia rocks |                | Marine | 29 |
| 600      | 12.69                  | Katanga oils | Oils/Bitumen | Marine | 28 |
| 625      | 5.4                    | Masirah Bay rocks | Oils/Bitumen | Marine | 27 |
| 720      | 8.8                    | Draken Fm. | Oils/Bitumen | Marine | 2 |
| 800      | 8.78                   | Baykit oils | Oils/Bitumen | Marine | 28 |
| 830      | 6.6                    | Bitter Springs | Oils/Bitumen | Marine | 2 |
| 850      | 18.7                   | Kwagunt Fm. | Oils/Bitumen | Marine | 2 |
| 1100     | 2                      | Tourist Fm. (Taoudeni Basin) |                | Marine | 30,31 |
| 1370     | 8.1                    | Xiamaling Fm. |                | Marine | 32 |
| 1380     | 8.5                    | Roper Group |                | Marine | 33 |
| 1400     | 5.8                    | Mc Minn Fm. | Oils/Bitumen | Marine | 2 |
| 1485     | 14.7                   | Yalco Fm. | Oils/Bitumen | Marine | 2 |
| 1600     | 9                      | Balbirini Fm. | Oils/Bitumen | Marine | 2 |
| 1640     | 8.5                    | Barney Creek Fm. | Oils/Bitumen | Marine | 2 |
| 1730     | 12                     | Wollogorang | Oils/Bitumen | Marine | 2 |

1. Kuypers et al., (2004); 2. Summons et al., (1999); 3. French et al., (2020); 4. Lyons et al., (2019); 5. Schaefer et al., (2020); 6. Sepúlveda et al., (2009); 7. Farrimond et al., (1990); 8. Dumitrescu and Brassell (2005); 9. Farrimond et al., (2004); 10. van Breugel et al., (2007); 11. Kujau et al., (2012); 12. Farrimond et al., (1989); 13. French et al., (2014); 14. Stüeken et al., (2019); 15. Kasprak et al., (2015); 16. McEvoy and Giger (1986); 17. Tian et al., (2013); 18. Salto et al., (2016); 19. Xie et al., (2005); 20. Cao et al., (2009); 21. Hays et al., (2012); 22. Haddad et al., (2016); 23. Rohrssen et al., (2012); 24. Pancost et al., (1998); 25. Spaak et al., (2017); 26. Lee et al., (2019); 27. Grosjean et al., (2009); 28. Kelly et al., (2011); 29. Lee et al., (2013); 30. Blumenberg et al., (2012); 31. Gueneli et al., (2018); 32. Luo et al., (2015); 33. Jarrett et al., (2019).
a higher frequency during the Mesozoic. Very few studies reported 2-MeHops during the Paleozoic, and, to the best of our knowledge, they have not been detected in Silurian rocks. For studies that report 2-MeHop, the maximum relative abundance (2-MeHops’ index) is <15% for most of Earth’s history. The few exceptions come from the Phanerozoic with samples from (i) Observatory Hills FM at ~535 Myr (Summons et al., 1999); (ii) Cominco SS-9 core at 455 Myr (Pancost et al., 1998); (iii) the Permian/Triassic mass extinction and its aftermath around 250 Myr (Cao et al., 2009; Saito et al., 2016); (iv) the end Triassic extinction around 201 Myr (Kasprak et al., 2015); (v) the Toarcian OAE, OAE 1a, and OAE 2 at 183, 120, and 94 Myr (Dumitrescu & Brassell, 2005; Farrimond et al., 1989); and (vi) the PETM at 56 Myr (Lyons et al., 2019). During these events, the 2-MeHops’ index reached values >40%.
4 | DISCUSSION

4.1 | 2-MeHops and hpnP gene in modern bacteria

Expanding on earlier work (e.g., Welander et al., 2010), our analysis shows that the hpnF/shc gene, responsible for the cyclisation of squalene into a hopanoid (Figure 1), is relatively widespread among a number of bacterial phyla (Figure 2). This is consistent with the detection of hopanoids in a wide range of bacteria (Rohmer et al., 1984). However, the hpnP gene, which is responsible for the methylation at the C2 position after the hopanoid skeleton is formed (Figure 1), is less widespread. It is predominantly found in cyanobacteria and α-proteobacteria (Figure 2), in accordance with other studies (Welander et al., 2010). One acidobacterium (Candidatus Koribacter versatilis) and three actinobacteria possess both the hpnF and hpnP genes and hence should be genetically capable of synthesizing 2-MeHops. Synthesis of 2-MeHops in this acidobacterium was recently confirmed by culture work (Sinninghe Damsté et al., 2017), although this appears to be an exception as most acidobacteria do not possess the hpnP gene and their cultures do not contain 2-MeHops. As Candidatus Koribacter versatilis is so far the only acidobacterium capable of 2-MeHops' synthesis, it is likely that this species obtained the hpnP gene through lateral gene transfer. Due to the low support values in this part of the gene tree, it is currently unclear whether the lateral gene transfer occurred from cyanobacteria or α-proteobacteria. So far 2-MeHops have not been found in cultures of actinobacteria, and future work should determine whether the specific actinobacteria with the hpnP gene produce 2-MeHop.

Work for species with available cultures confirms our genomic analyses, as cultures of bacteria that possess the hpnP gene produce 2-MeHops and those that lack the hpnP gene do not (Figure 2). However, many species have not yet been analyzed for their 2-MeBHP content in culture and/or have not had their full genome analyzed, so we cannot generalize our findings yet. In addition, there is a small number of discrepancies (highlighted in Figure 2 with arrows 1 to 4). Methylcystis parvus and Leptolyngbya boryana have the
hpnF and hpnP genes, but the original culture work does not report the occurrence of 2-MeBHPs (Rohmer et al., 1984). It is likely that these bacteria did not produce significant amounts of 2-MeBHPs under the specific culture conditions as the amount of 2-MeBHP varies depending on culture conditions (Doughty et al., 2009; Wu et al., 2015).

More problematic is the reported production of 2-MeHops by Halothece sp. PCC 7418 (originally reported as Synechococcus ATCC 29534) and Synechococcus elongatus PCC 6301 (reported as ATCC 27144) (Summons et al., 1999), as these species lack both the hpnF and hpnP genes. To explore these discrepancies further, we analyzed biomass from Synechococcus elongatus PCC 7942 for the presence of (2-methyl)hopanoids. The gene organization and sequence of Synechococcus elongatus PCC 7942 is nearly identical to Synechococcus elongatus PCC 6301 as their 16S and 18S ribosomal RNA sequences are 100% identical, and out of the 136 protein sequences that we tested, 100 have 100% identity, 35 are more than 99% similar, and the remaining protein is 98.5% similar for the two strains. However, unlike PCC 6301, PCC 7942 is available as culture. Our HT-GC-MS analysis (Figure 3) shows no compounds in the m/z 191 and 205 ion traces, demonstrating that this cyanobacterium does not make hopanoids, consistent with the lack of hpnF and hpnP genes in its genome. Thus, we conclude that the earlier report of 2-MeBHP production by Synechococcus elongatus PCC 6301 (originally ATCC27144) was likely incorrect. The same might hold for Halothece sp. PCC 7418. Altogether, the culture work is therefore possibly consistent with the genomic analyses.

4.2 | 2-MeHops in the geological record

It is important to note that the 2-MeHop index in the rock record might reflect local/regional perturbations and not necessarily a global (marine) regime. This especially applies to the Proterozoic and early Phanerozoic where sample density is low and from which deep marine sediments are not preserved. 2-MeHops are ancient lipids observed in samples as old as the paleo-Proterozoic with samples from the Wollogorang Fm (1730 Myr) currently the oldest robust finding (Figure 4). Robust estimates of the 2-MeHops’ index thoroughly checked for contamination using the latest methods are yet not available for the Paleoproterozoic but for some sections from the Mesoproterozoic: the 1380 Myr old Rober Group (Jarrett
et al., 2019) and 1100 Myr old Taouden Basin (Blumenberg et al., 2012; Gueneli et al., 2018) and these studies confirm the occurrence of 2-methylhopanoids and hence bacterial life during the Mesoproterozoic.

For the Phanerozoic, while most reports come from marine sediments, 2-MeHops are also present in lacustrine sediments (e.g. Farrimond et al., 2004; French et al., 2020) and marine oils and bitumens (Summons et al., 1999). They are more frequently reported from Mesozoic samples, likely due to a higher sampling frequency compared with the Paleozoic and Precambrian. The highest reported 2-MeHops’ index in Earth’s history (maximum of 75%) is from samples from the marginal South Dover Bridge site in North America deposited during the Paleocene Eocene Thermal Maximum (PETM) at 56 Myr (Lyons et al., 2019). However, this site received significantly amounts of reworked organic matter from organic-rich Cenomanian rocks during the PETM (Lyons et al., 2019). Hence, the reported high 2-MeHops might not reflect in situ production during the PETM but rather reflects production during the Cenomanian, where other studies report a high 2-MeHops’ index (Figure 4).

2-MeHops become particularly abundant (>15%) during specific events characterized by major climatic, biotic, and environmental changes in the Phanerozoic. Almost all these events are characterized by major changes in ocean biogeochemistry and anoxia. Highest relative abundance of 2-MeHops’ indices occurs under low oxygen/anoxic conditions, but when de-oxygenation intensifies and euxinic conditions develop, the relative abundance of 2-MeHops decreases. This dependency on the extent of anoxia is especially clear in the Ordovician Cominco SS-9 core (~455 Myr). In that record, 2-MeHops’ relative abundance is highest in the Platville Fm (max. 2-MeHops’ index of 21%), deposited under anoxic conditions with TOC levels of 0.5–2 wt. % (Pancost et al., 1998). When anoxia intensified and TOC levels increased to >40 wt. % in the Decorah FM and euxinic conditions reached the photic zone, the 2-MeHops’ index rapidly decreased to values around 6%. Similarly, 2-MeHops are absent from Demarara Rise during OAE 2, when anoxia was intense (TOC >40 wt. %) and biomarker evidence indicates that euxinic conditions episodically reached the photic zone (Forster et al., 2004). The high 2-MeHops’ relative abundance observed in the Phanerozoic in samples deposited under low oxygen conditions and perturbed marine biogeochemical cycles (e.g., N cycle) agrees with Ricci et al.’s (2013) findings, showing that the hpnP gene is enriched in the modern settings when the environment is poor in oxygen and fixed nitrogen.

4.3 | What drove the elevated sedimentary occurrence of 2-MeHop?

Most reports of 2-MeHops come from marine sediments spanning specific events (Kuypers et al., 2004; Xie et al., 2005; Figure 4). What drove this episodic relative increase in 2-MeHop? Based on our compilation of the distribution of the hpnP gene in bacterial genomes and 2-MeBHPs in cultures (expanding on earlier work like (Welander et al., 2010)), we saw that at least three marine strains can synthesize 2-MeBHPs, the cyanobacterium filamentous cyanobacterium ESFC-1 and the α-proteobacteria Methylobacterium salsuginis and Nitrobacter sp. Nb-311A (Figure 5). The actinobacteria and the acidobacterium that possess the hpnP gene are terrestrial strains. Information found in bacterial genomes can be used to infer their evolutionary history and ancestral habitats (Sánchez-Baracaldo et al., 2021), and we recognize, however, that modern conditions might not be indicative of ancestral environmental conditions and modern habitats do not always provide information about the complete environmental (e.g., salinity) tolerance species possess.

Of the three marine strains, filamentous cyanobacterium ESFC-1 has recently been identified as an aerobic nitrogen-fixing cyanobacterium (Eversroad et al., 2016; Woebken et al., 2012). To our knowledge, this is the first and only marine N₂-fixing cyanobacterium holding the hpnP gene. An expansion of (ancestors of) this species could fulfill the original hypothesis (Kuypers et al., 2004) that the observed elevated sedimentary occurrence of 2-MeHops during OAEs (Figure 4) is driven by an increase in marine N₂-fixation in response to major marine N-cycle perturbations within OAEs (Naafs et al., 2019). However, so far filamentous cyanobacterium ESFC-1 has only been identified in an intertidal coastal microbial mat (Woebken et al., 2012) and has not been reported in open marine environments. Microbial mats could have contributed to 2-MeHops’ sedimentary record in Proterozoic samples as microbial mats were much more widespread before the Cambrian revolution (Gehling, 1999). However, it seems unlikely that 2-MeHops produced in coastal microbial mats dominated the hopanoid inventory in open ocean sediments during more recent events of the Phanerozoic such as the OAEs.

The occurrence of 2-MeHops during the Phanerozoic (Figure 4) might therefore reflect an expansion of marine α-proteobacteria. Various studies proposed that α-proteobacteria produce more 2-MeBHPs under hypoxic and acidic conditions (Kulkarni et al., 2015; Wu et al., 2015). This observation agrees with the observed 2-MeHops’ occurrence in the geological record, in which the highest relative abundance is associated with widespread (water column) hypoxic/anoxic conditions such as during the OAEs. However, so far, all studies on the impact of cell stress and 2-MeBHPs come from plant symbionts that operate in the terrestrial realm. It is not clear whether these results can be extrapolated to the marine realm.

Out of all modern α-proteobacteria that contain the hpnP gene, only two species have a marine habitat: Methylobacterium salsuginis and Nitrobacter Nb-311A (Figure 5). Methylobacterium salsuginis is strictly aerobic and facultatively methylotrophic (Wang et al., 2007). Other Methylobacteria produce 2-MeBHPs in culture (Knani et al., 1994) and many have the hpnP gene (Figure 5), but Methylobacterium salsuginis is the only known marine strain holding the hpnP gene. This Methylobacterium reduces nitrate into nitrite (Green & Ardley, 2018), which is a fundamental step of denitrification, a main sink of marine nitrogen. Biogeochemical modeling studies indicate that denitrification rates were high during OAEs (Naafs et al., 2019),
suggesting that high observed 2-MeHops indicates the spread of denitrifiers. The high 2-MeHops’ relative abundance observed during some of the geological events associated with widespread marine anoxia (Figure 4) could therefore be (partly) explained by an expansion of ancestors of Methylobacterium salsuginis and reflect high rates of denitrification.

Nitrobacter Nb-311A is the second marine α-proteobacterium to contain the hpnP gene (Welander et al., 2010). This bacterium is strictly aerobic and oxidizes nitrite to nitrate during nitrification (the opposite process of denitrification), a crucial process in the marine N cycle. As with denitrification, biogeochemical modeling indicates that nitrification rates increased by one order of magnitude during OAEs (Naafs et al., 2019), which would also agree with the observed high 2-MeHops’ index. The high 2-MeHops’ index during some of the geological events (Figure 4) could therefore (partly) be explained by an expansion of Nitrobacter that drove the high rates of nitrification. A similar hypothesis was recently proposed by Elling et al., (2020). They used cultures of Nitrobacter vulgaris AB1 and argued the high 2-MeHops’ level observed during OAEs reflects an expansion of Nitrobacter in response to high rates of nitrification. However, Nitrobacter vulgaris AB1 has so far only been found in terrestrial settings (Mellby et al., 2017; Vanparys et al., 2007). Besides, the marine habitat of Nitrobacter Nb-311A is not well constrained, based solely on work in 1968 and a personal communication in Starkenburg et al., (2008). The Starkenburg paper also doubted the assumed marine habitat of Nitrobacter Nb-311A as its SSU rRNA gene sequence is identical to the one of Nitrobacter winogradskyi, an organism only found in soil (Poly et al., 2008). Thus, while there is some evidence that 2-MeHops’ occurrence indicates high rates of denitrification in the ocean (via Methylobacterium salsuginis), the marine habitat of Nitrobacter spp. is not well established and requires further research.

5 | CONCLUSIONS

2-Methylhopanes are among the oldest biomarkers on Earth, but the ecological and environmental interpretations of their occurrence in marine sediments throughout the Phanerozoic and Precambrian are poorly constrained. By combining culture data with a broad genomic sampling, we demonstrated that 2-MeBHPs are almost exclusively produced by cyanobacteria and α-proteobacteria, with only two open ocean species currently known. Both open marine species are α-proteobacteria, which are involved in biogeochemical processes fundamental to the marine N-cycle’s operation. Although preferred habitats of modern organisms might not necessarily inform us about their ancestral environmental tolerances in the geological past, particularly since paleoenvironments and biogeochemical cycles were different in the past, we propose that the episodic occurrence of high 2-MeHops’ levels as observed during the Phanerozoic might be driven by an expansion of marine α-proteobacteria in response to a perturbation of the marine N cycle.

ACKNOWLEDGEMENTS

BDAN was funded through a Royal Society Tata University Research Fellowship, GB was funded by a postgraduate scholarship from the University of Bristol, and PS-B was funded through a Royal Society University Research Fellowship, while FMM was supported by a NERC standard grant (NE/N011112/1). We thank Simon Brassell, Roger Summons, Jochen Brocks, Shelby Lyons, and Kate French for their help with compiling the 2-MeHops’ data from the geological record. Felix Elling is acknowledged for discussing the modern 2-MeBHP producers, and we thank Jenan Kharbush for providing the Synechococcus biomass. We thank Ian Bull for help with the HT-GC-MS system. Bioinformatic analyses were carried out using the computational facilities of the Advanced Computing Research Centre, University of Bristol—http://www.bris.ac.uk/acrc/. The data that support the findings of this study are available in the supplementary material of this article.

ORCID

B. D. A. Naafs https://orcid.org/0000-0001-5125-6928

REFERENCES

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2017). GenBank. Nucleic Acids Research, D37–D42. https://doi.org/10.1093/nar/gkw1070

Bisseret, P., Zundel, M., & Rohmer, M. (1985). Prokaryotic tri-terpenoids: 2,3β-methylhphanoids from Methylobacteriumorganophilum and Nostoc muscorum, a new series of prokaryotic triterpenoids. European Journal of Biochemistry, 150, 29–34. https://doi.org/10.1111/j.1432-1033.1985.tb08982.x

Bligh, E. G., & Dyer, W. J. (1959). A rapid method of total lipid extraction and purification. Canadian Journal of Biochemistry and Physiology, 37, 911–917. https://doi.org/10.1139/o59-099

Blumenberg, M., Thiel, V., Riegel, W., Kah, L. C., & Reitrner, J. (2012). Biomarkers of black shales formed by microbial mats, Late Mesoproterozoic (1.1Ga) Taoudeni Basin, Mauritania. Precambrian Research, 196–197, 113–127. https://doi.org/10.1016/j.prec res.2011.11.010

Bravo, J.-M., Perzl, M., Härtter, T., Kannenberg, E. L., & Rohmer, M. (2001). Novel methylated triterpenoids of the gammacerane series from the nitrogen-fixing bacterium Bradyrhizobium japonicum USDA 110. European Journal of Biochemistry, 268, 1323–1331. https://doi.org/10.1046/j.1432-1377.2001.01998.x

Brocks, J. J., Logan, G. A., Buick, R., & Summons, R. E. (1999). Archean molecular fossils and the early rise of euukaryotes. Science, 285, 1033–1036. https://doi.org/10.1126/science.285.5430.1033

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: architecture and applications. BMC Bioinformatics, 10, 421. https://doi.org/10.1186/1471-2105-10-421

Cao, C., Love, G. D., Hays, L. E., Wang, W., Shen, S., & Summons, R. E. (2009). Biogeochemical evidence for euxinic oceans and ecological disturbance presaging the end-Perman mass extinction event. Earth and Planetary Science Letters, 281, 188–201. https://doi.org/10.1016/j.epsl.2009.02.012

Cvejic, J. H., Putra, S. R., El-Beltagy, A., Hattori, R., Hattori, T., & Rohmer, M. (2000). Bacterial triterpenoids of the hopane series as biomarkers for the chemotaxonomy of Burkholderia, Pseudomonas and Ralstonia spp. FEMS Microbiology Letters, 183, 295–299. https://doi.org/10.1111/j.1574-6968.2000.tb08974.x
carbon into the atmosphere. *Paleoceanography*, 22, PA1210. https://doi.org/10.1029/2006pa001341

Vanparys, B., Spieck, E., Heylen, K., Wittebolle, L., Geets, J., Boon, N., & De Vos, P. (2007). The phylogeny of the genus *Nitrobacter* based on comparative rep-PCR, 16S rRNA and nitrite oxidoreductase gene sequence analysis. *Systematic and Applied Microbiology*, 30, 297–308. https://doi.org/10.1016/j.syapm.2006.11.006

Vanparys, B., Spieck, E., Heylen, K., Wittebolle, L., Geets, J., Boon, N., & De Vos, P. (2007). The phylogeny of the genus *Nitrobacter* based on comparative rep-PCR, 16S rRNA and nitrite oxidoreductase gene sequence analysis. *Systematic and Applied Microbiology*, 30, 297–308. https://doi.org/10.1016/j.syapm.2006.11.006

Vilcheze, C., Llopiz, P., Neunlist, S., Poralla, K., & Rohmer, M. (1994). Prokaryotic triterpenoids: new hopanoids from the nitrogen-fixing bacteria *Azotobacter vinelandii*, *Beijerinckia indica* and *Beijerinckia mobilis*. *Microbiology*, 140, 2749–2753. https://doi.org/10.1099/00221287-140-10-2749

Waldbauer, J. R., Sherman, L. S., Sumner, D. Y., & Summons, R. E. (2009). Late Archean molecular fossils from the Transvaal Supergroup record the antiquity of microbial diversity and aerobiosis. *Precambrian Research*, 169, 28–47. https://doi.org/10.1016/j.precamres.2008.10.011

Wang, X., Sahr, F., Xue, T., & Sun, B. (2007). *Methylobacterium salsuginis* sp. nov., isolated from seawater. *International Journal of Systematic and Evolutionary Microbiology*, 57, 1699–1703. https://doi.org/10.1099/ijsem.0.64877-0

Welander, P. V., Coleman, M. L., Sessions, A. L., Summons, R. E., & Newman, D. K. (2010). Identification of a methylase required for 2-methylhopanoid production and implications for the interpretation of sedimentary hopanes. *Proceedings of the National Academy of Sciences*, 107, 8537–8542. https://doi.org/10.1073/pnas.0912949107

Woebken, D., Burow, L. C., Prufert-Bebout, L., Bebout, B. M., Hoehler, T. M., Pett-Ridge, J., Spormann, A. M., Weber, P. K., & Singer, S. W. (2012). Identification of a novel cyanobacterial group as active diazotrophs in a coastal microbial mat using NanoSIMS analysis. *The ISME Journal*, 6, 1427–1439. https://doi.org/10.1038/is mej.2011.200

Wu, C.-H., Bialecka-Fornal, M., & Newman, D. K. (2015). Methylation at the C-2 position of hopanoids increases rigidity in native bacterial membranes. *eLife*, 4, e05663. https://doi.org/10.7554/eLife.05663

Xie, S., Pancost, R. D., Yin, H., Wang, H., & Evershed, R. P. (2005). Two episodes of microbial change coupled with Permo/Triassic faunal mass extinction. *Nature*, 434, 494–497. https://doi.org/10.1038/nature03396

Zhao, N., Berova, N., Nakanishi, K., Rohmer, M., Mougenot, P., & Jürgens, U. J. (1996). Structures of two bacteriohopanoids with acyclic pentol side-chains from the cyanobacterium Nostoc PCC 6720. *Tetrahedron*, 52, 2777–2788. https://doi.org/10.1016/0040-4020(96)00013-0

Zundel, M., & Rohmer, M. (1985). Prokaryotic triterpenoids 1. 3β-Methylhopanoids from *Acetobacter* species and *Methylococcus capsulatus*. *European Journal of Biochemistry*, 150, 23–27. https://doi.org/10.1111/j.1432-1033.1985.tb08980.x

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section.

---

**How to cite this article:** Naafs, B. D. A., Bianchini, G., Monteiro, F. M., & Sánchez-Baracaldo, P. (2022). The occurrence of 2-methylhopanoids in modern bacteria and the geological record. *Geobiology*, 20, 41–59. [https://doi.org/10.1111/gbi.12465](https://doi.org/10.1111/gbi.12465)