Cellular and Molecular Biological Approaches to Interpreting Ancient Biomarkers

Dianne K. Newman,[1,2,3] Cajetan Neubauer,[1,2,3] Jessica N. Ricci,[1] Chia-Hung Wu,[1] and Ann Pearson[4]

[1]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, California 91125; email: dkn@caltech.edu

[2]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125

[3]Howard Hughes Medical Institute, Pasadena, California 91125

[4]Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts 02138

Annu. Rev. Earth Planet. Sci. 2016. 44:493–522

First published online as a Review in Advance on May 9, 2016

The *Annual Review of Earth and Planetary Sciences* is online at earth.annualreviews.org

This article's doi: 10.1146/annurev-earth-050212-123958

Keywords

biomarkers, molecular fossils, interpretation, hopanoids, evolution, rock record

Abstract

Our ability to read the molecular fossil record has advanced significantly in the past decade. Improvements in biomarker sampling and quantification methods, expansion of molecular sequence databases, and the application of genetic and cellular biological tools to problems in biomarker research have enabled much of this progress. By way of example, we review how attempts to understand the biological function of 2-methylhopanoids in modern bacteria have changed our interpretation of what their molecular fossils tell us about the early history of life. They were once thought to be biomarkers of cyanobacteria and hence the evolution of oxygenic photosynthesis, but we now believe that 2-methylhopanoid biosynthetic capacity originated in the Alphaproteobacteria, that 2-methylhopanoids are regulated in response to stress, and that hopanoid 2-methylation enhances membrane rigidity. We present a new interpretation of 2-methylhopanes that bridges the gap between studies of the functions of 2-methylhopanoids and their patterns of occurrence in the rock record.

1. INTRODUCTION

Understanding the history of life on Earth is a profound challenge. When did particular organisms first appear? When did they flourish or go extinct? With whom did they associate? Where did they live? What combination of physical, chemical, and biological changes drove evolutionary and ecological dynamics? Such questions have beguiled scholars from the beginning of recorded history, and many attempts have been made to answer them, from the *Book of Genesis* and other creation myths to *On the Origin of Species* to modern theories of evolutionary developmental biology (Carroll 2008). Yet deciphering life's origins and evolution is hard, if not fundamentally impossible, particularly the farther back in time we go. Earth did the experiment once, but we cannot know precisely what it did, nor can we repeat it.

Despite these challenges, or perhaps because of them, understanding the past holds a powerful intellectual allure. Fortunately, technological and conceptual advances have made the construction of meaningful narratives about life's history more possible today than ever before. Access to diverse records, from geochemical to genomic, has expanded greatly in the past decade. Not only have more samples been taken of ancient rocks, shared internationally, and analyzed by standardized geochemical protocols (e.g., French et al. 2015), but innovations in DNA sequencing have resulted in exponentially growing genomic databases, permitting insights into evolution through comparative genomics (David & Alm 2011). Advances in microanalysis have made it possible to identify patterns on spatial scales that were unresolvable until recently, leading to revisions in the interpretation of ancient evolutionary events (Rasmussen et al. 2008, Williford et al. 2013, Fischer et al. 2014).

Alongside these technical advances has come a less tangible but no less important form of progress: the rise of a new generation of geobiologists who are trained in molecular and cellular biology. Our goal in this review is to explain why this has been so helpful to the study of evolution in deep time. Microorganisms dominated the biosphere throughout Earth history, and unlike the dinosaurs, they left behind a frustratingly ambiguous morphological fossil record. With a few notable exceptions (e.g., Petroff et al. 2010 and Crosby et al. 2014), very little information can be extracted from microbial shape—be it at the single-cell or community level—to infer the phylogenetic composition and/or metabolic properties of ancient microorganisms. If we seek insight into evolutionary processes early in the history of life, we must therefore embrace the molecular fossil record. And a thoughtful interpretation of this record requires knowledge of molecular and cellular biology.

To support this contention, we begin with a discussion of the challenges inherent in studying diverse types of molecular fossils. From there, we provide a brief background on lipid biology to orient the geoscience reader to well-established biological concepts that are relevant to our discussion. We narrow our focus to a particular class of biomarkers: hopanes, which are sedimentary remnants of hopanoids, believed to be steroid surrogates in bacteria. Our intent is not to provide a comprehensive review of hopanoid geochemistry or biological function; rather, we use the story of hopanoids—and one hopanoid subtype in particular, 2-methylhopanoids—to illustrate how insights and approaches from modern molecular and cellular biology can add to the understanding of ancient life.

2. FOSSIL TYPES, TEMPORAL RANGE, AND INTERPRETATIVE CHALLENGES

Fossils, the preserved remains and traces of ancient organisms, can be used to unravel the history of life on Earth. How does the molecular fossil record compare to the morphological fossil record or to the evolutionary histories retained within modern genomes (**Figure 1*a***)? Molecular fossils

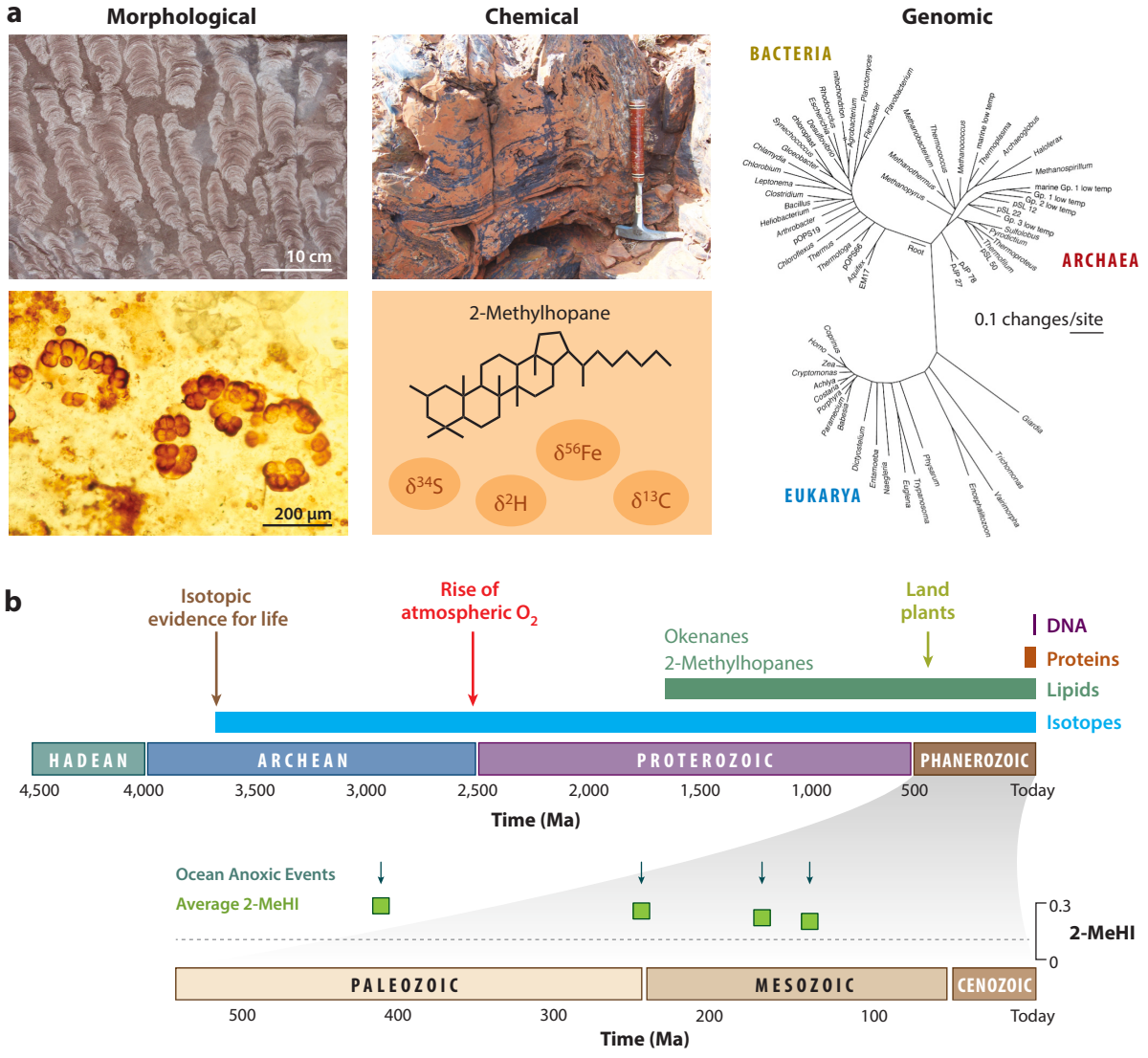


Figure 1

Geological and biological imprints of ancient life. (*a*) Three types of evolutionary records. Morphological: small Mesoproterozoic columnar stromatolites from the Southern Urals, Russia (*top*; photo credit: Tanja Bosak); population of silicified *Eoentophysalis*, interpreted as colonial coccoidal (and mat-forming) cyanobacteria, from the 1,200–1,300 Ma Debengda Formation, northern Siberia (*bottom*; photo credit: Andrew Knoll). Chemical: banded iron formation from the 2.5–2.2 Ga Gamohaan Hill site, South Africa (*top*; photo credit: Andreas Kappler); 2-methylhopane structure and common isotope-based proxies (*bottom*). Genomic: phylogenetic tree (Pace 1997). (*b*) Selected geobiological data in the context of geological time. The oldest known evidence for biomolecules in the rock record is shown (Willerslev et al. 2003, Brocks et al. 2005, Schweitzer et al. 2009). Phanerozoic 2-methylhopane indices (2-MeHI) measured during Oceanic Anoxic Events are displayed; the average 2-methylhopane index across this time frame is denoted by the gray dashed line (data adapted from figure 5 in Knoll et al. 2007). 2-MeHI = C_{30} 2-MeHI/(C_{30} 2-MeHI + C_{30} desmethyl hopane).

comprise a vast number of organic and inorganic compounds and can be defined as molecules with a characteristic chemical structure that are remnants of diverse biomolecules that once constituted living cells (e.g., DNA, protein, lipids). In contrast, morphological fossils are structures whose interpretation is based upon their shape, be it at the nanoscale (e.g., magnetite particles), microscale (e.g., cell bodies), millimeter scale (e.g., roll-up mat-like structures), or meter scale (e.g., extensive reef-like forms). Some morphological fossils are the physical remains of organisms that have mineralized, whereas others are imprints, traces, or burrows that life has left behind. The morphology of these fossil types can be compared to modern species to decipher their evolutionary relationships. Though morphological fossils can be found in Precambrian rocks, their interpretation is often plagued by ambiguity (Brasier et al. 2002, Wacey et al. 2011, Cunningham et al. 2012). Even if consensus can be reached that a particular structure derived from a biological rather than an abiotic process (Allwood et al. 2006, Schopf et al. 2015), it can be challenging to infer much about the nature of these organisms from morphology alone. This is because diverse microbes often look alike. More often than not, we need additional information from molecular sources to gain insight into bacterial physiology.

Molecular fossils come in many forms and are stable over different periods of time (**Figure 1*b***). Organisms can leave molecular fossils by disturbing isotope records (e.g., preferentially enriching specific isotopes into biomass as a result of enzyme activity) or by depositing organic structures that can be preserved as biomarkers. Biomarkers carry information about their source organism in their sequence of nucleotides (DNA or RNA) or amino acids (protein) or within their lipid structures. Unfortunately, with the exception of a few rare cases, DNA, RNA, and protein degrade quickly on geologic timescales (Brocks & Banfield 2009). Due to rapid hydrolysis, RNA is an extremely short-lived molecule with a half-life on the timescale of hours. DNA and protein have been recovered from significantly older samples, but these instances are limited to particular—often sulfidic—environments. For example, 600,000-year-old plant DNA has been retrieved from reducing sediments; in oxidizing environments, the half-life of DNA is typically on the order of weeks (Willerslev et al. 2003, Pedersen et al. 2015). Additionally, collagen protein sequences have been recovered from dinosaur bones that date to 80 to 68 Ma (Asara et al. 2007, Schweitzer et al. 2009, Bertazzo et al. 2015), yet no other proteins were found.

Unique among molecular fossils, lipids are preserved for extended lengths of time (Peters et al. 2005). Indeed, lipid biomarkers provide a view into Earth history that is at least as long as that provided by morphological fossils. The lipid biomarker field is vast, and we refer the reader to two excellent reviews that document its history and key discoveries (Brocks & Summons 2003, Whiteside & Grice 2016). Lipid molecules that survive diagenesis, the process of sediment becoming rock, may be chemically altered (e.g., isomerization or reduction) before they become biomarkers, but typically the hydrocarbon skeleton remains recognizable as derived from—if not identical to—the parent molecule. This remarkable property of lipids manifests in the form of extractable petroleum deposits, the oldest of which date to 1.64 Ga in the Mesoproterozoic and are undisputedly the remnants of ancient ecosystems (Brocks et al. 2005). In many places, lipid biomarker concentrations are far too low for economic viability, yet remain well above the threshold needed for biogeochemical analyses; it is from these rocks that we gain much of our insight into life on the ancient Earth. Lipid fossils have helped establish the existence of multicellular animals prior to 635 Ma (Love et al. 2009) and the presence of the oxygen-dependent sterol biosynthetic pathway by at least 1,640 Ma (Brocks et al. 2005). However, no syngenetic lipids have yet been found in Archean rocks (Brocks et al. 1999, Rasmussen et al. 2008, Brocks 2011, Pawlowska et al. 2013, French et al. 2015).

Fossils, be they molecular or morphological, permit us to connect ancient life forms to discrete moments in time. However, even assuming accurate fossil dating, it is challenging to infer the

temporal order of evolutionary events from the traditional fossil record. This is because the record is incomplete. We have limited access to ancient rocks for two reasons: (*a*) inherent lack of availability (many rocks deposited early in Earth history have since been destroyed or transformed through tectonic processes, and not all organisms leave recognizable fossils), and (*b*) practical challenges in accessing the rock record (some of the best-preserved ancient deposits are found in relatively inaccessible parts of the globe). Accordingly, though morphological and molecular fossils can help constrain the evolutionary timeline, we must turn to other records for complementary information.

Sequence information, as captured in nucleotides or amino acids, provides a rich record of evolutionary history. If we know what to look for—such as a gene or protein that is diagnostic of a group of organisms or a biological process of interest—we can infer evolutionary relationships between multiple sequences for that gene or protein and others. This is called phylogeny, the science of using sequence information to infer evolutionary relationships. For example, Carl Woese pioneered the use of ribosomal RNA as a molecular chronometer to achieve a coherent organismal classification scheme that defined the three domains of life (Woese & Fox 1977). As long as a sufficient number of sequences are available to compare to one another, it is possible to construct phylogenetic trees to infer the sequence of evolutionary events. Confidence in phylogenetic assignments also depends on the algorithms used to infer the tree, and as bioinformatics becomes more central to modern biology, these algorithms are continually improving (Pais et al. 2014). We emphasize that rooted phylogenetic trees provide information about the relative temporal order of evolutionary events; they can be used to date processes only if they are calibrated by fossils that constrain specific branches on the tree. Even under the best circumstances (i.e., an extensive fossil record that can help date multiple branches on the tree), constructing reliable molecular clocks is challenging because rates of molecular evolution are not constant and it is not always clear where to place fossils (Bell 2015). Accordingly, it usually is best to refrain from making claims about absolute dates based on phylogenetic analysis. Moreover, knowing that something was present at a specific moment in time, or even that it appeared prior to or after something else, does not tell us why it appeared when it did. If we seek to understand what triggered key biological developments (e.g., origin of life, origin of oxygenic photosynthesis, multicellularity), we must go beyond correlations and enter the challenging realm of causality. And to interpret causality, we must bridge the gap between molecular structure and biological function.

3. LIPIDS IN THE CELL

Robust interpretation of a biosignature requires awareness and understanding of all the different mechanisms that may have led to its creation. For example, when we see a fossilized microbial mat in the rock record, it is important to recognize that microbial communities form mat-like structures in diverse environments and can be composed of very different organisms. If our goal were to infer the type of microbial community from a fossilized mat with respect to which specific organisms and/or metabolisms were once active in that mat, we would need to know what to look for to differentiate among various organisms. Whether this is achievable depends on the organisms or processes being compared, and whether they leave specific and preservable biosignatures. Moreover, we would need to consider how the preservation process itself impacts the material being fossilized (e.g., Pawlowska et al. 2013). Similarly, to properly interpret molecular biomarkers, we must appreciate the range of roles particular molecular types can play in living cells. We assume uniformitarian principles apply at the biochemical level: that molecular functions in living cells today are likely to be the same as (or very similar to) their roles in ancient cells. To set the stage, we briefly summarize the primary biological functions of lipids in the remainder of this section.

3.1. Diversity of Lipid Structures

Cellular lipids serve many functions; for example, they can be essential structural components of membranes, storage compounds for energy and carbon, or signaling molecules (Wenk 2010). Like other metabolites, lipids are not directly encoded in the genome, but are generated via enzymes whose activity can be altered by many variables. Organisms make a marvelous variety of lipid types. Biological extracts such as from blood serum may contain hundreds of thousands of distinct lipids (Baker et al. 2014). This complexity results from combinatorial innovations based upon a few common building blocks.

Over the past decade, the lipid research community, organized by the LIPID MAPS consortium, has developed guidelines for the classification of biological lipids (Fahy et al. 2009). The most common lipids are based on a glycerol backbone linked to hydrocarbon tails. In bacteria and eukaryotes, the chains are usually carboxylic acids that are produced by rounds of condensation and reduction of acetyl coenzyme A. Glycerophospholipids and glycosylglycerols additionally contain a polar head group and are typical bulk components of membranes. In most archaeal lipids, the chains are polymers of isoprene that are linked to glycerol via ether bonds and a distinct stereochemistry. Lipids based on a backbone of sphingosine are less abundant than glycerolipids but can play key cellular roles in structuring the membrane and as signaling molecules that mediate responses to stresses (Breslow & Weissman 2010). In saccharolipids, fatty acyls are linked directly to a sugar backbone. The Lipid A from lipopolysaccharide in gram-negative bacteria belongs to this category. Polyprenols (e.g., hopanoids, carotenoids, and sterols) are all polymers of isoprene. Lipid biosynthetic pathways and the phylogenetic distribution of biosynthetic genes of geological biomarkers have been reviewed recently (Pearson 2014).

Even after their synthesis, many lipids can undergo further modification, such as cyclization, saturation/desaturation, methylation, deacetylation, and addition of polar functional groups. These structural modifications can regulate physiological function because function is strictly linked to chemical structure.

3.2. Chemical Properties of Lipids

Lipids are hydrophobic or amphipathic molecules that tend to aggregate in water. The release of solvent molecules from aggregation increases entropy and also results in a gain of enthalpy through greater interaction between the released solvent molecules (Tanford 1978). This hydrophobic effect is strengthened by weak attraction forces that act at short range between nonpolar molecules. These basic processes give rise to a wide range of emergent properties of membranes that can be measured experimentally but are hard to predict theoretically due to the large number of small energy differences involved (Sackmann 1995).

Lipids in water form micelles (loosely bound aggregates with a hydrophobic core) or liposomes (globular lipid bilayers enclosing water, similar to cellular vesicles). The phase transitions of lipid mixtures and biological membranes can be measured in liposomes. When membranes contain longer and more saturated fatty tails, the contact area in the hydrophobic part of the membrane increases, and so do the interactions between the hydrocarbon chains. As a result, longer and more saturated tails typically increase melting temperature and decrease the fluidity of membranes. In analogy, lipid packing also impacts membrane viscosity, because lipids with good shape complementarity have more contact area. At the water-membrane interface, the hydrophilic moieties of lipids affect membrane stability, curvature, and charge. For any discussion of physiological roles of lipids it is necessary to understand that biological membranes contain a large proportion of proteins, typically 50–75% (w/w). Most cells regulate the composition of proteins and lipids to achieve a liquid-ordered state in which lipids and proteins can move laterally across the

membrane. In addition, lipids and proteins interact to orchestrate lateral compartmentalization within the membrane (lipid rafts) (Lingwood & Simons 2010).

3.3. Physiological Roles

Membranes separate individual cells from their environments and compartmentalize the cell interior into structures that carry out special tasks. This compartmentalizing function is so central that lipids must have been essential since the origin of life when proliferation of primitive cells could have occurred purely by biophysical processes (Lombard et al. 2012, Mercier et al. 2012, Blain & Szostak 2014).

The physiological functions that emerge for specific lipids are often context dependent. Cells have distinct membrane compartments that differ in their protein and lipid composition. Additionally, the two leaflets of a lipid bilayer are often asymmetric. Perhaps the best-studied lipid types with regard to subcellular distribution and biological functions are sterols, which can be used to illustrate that structurally related lipids can play many different functional roles (**Figure 2**). The concentration of sterols increases during the secretory pathway in eukaryotes, being low in the endoplasmic reticulum, higher in the Golgi, and highest in the plasma membrane (van Meer et al. 2008). Along the secretory pathway, membrane thickness increases, and proteins with longer hydrophobic sequences enrich jointly with sterols (Sharpe et al. 2010). At the plasma membrane, sterols, phosphosphingolipids, and proteins interact to form a mechanically robust membrane and promote membrane heterogeneity (Simons & Ikonen 1997, Hancock 2006, Lingwood & Simons 2010). In vitro, structurally different sterols have been shown to modulate membrane curvature, in some cases generating convex shapes and in others concave (Bacia et al. 2005).

The ability of microorganisms to control the biophysical properties of their membranes allows them to thrive in a wide range of environments. Because the hydrocarbon components of the membrane are the most energetically expensive modules to produce and have a major impact on

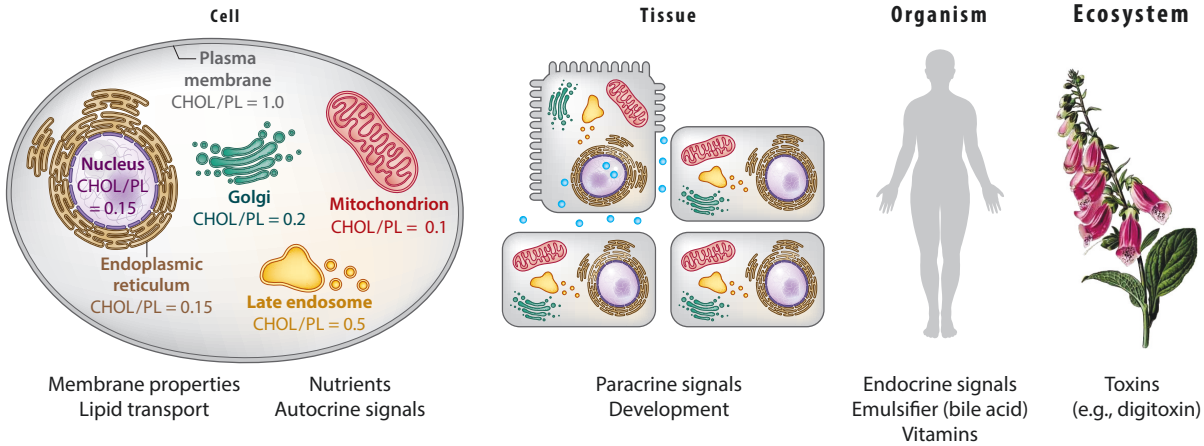


Figure 2

Summary of cellular distribution and biological functions of steroids. Steroids are structurally related to hopanoids and functionally better understood. In mammalian cells the ratio of cholesterol to phospholipids (CHOL/PL) generally increases during the secretory pathway from the endoplasmic reticulum to the Golgi and is highest in the plasma membrane. Apart from affecting the biophysical properties of membranes, steroids have biological functions that range from acting within individual cells to promoting interspecies interactions.

membrane fluidity, the regulation of fatty acid production is very tightly coupled to the growth condition. For example, cells can sense changes in the physical properties of the membrane and reconfigure the structure of existing phospholipids in response (Puth et al. 2015); they can also adapt by synthesizing new lipids and degrading others (Zhang & Rock 2009, Cybulski et al. 2015). Finally, the cellular lipid content can change in response to extracellular signals, metabolic and developmental states, and available nutrients. For instance, many cyanobacteria and algae can replace phospholipids with nonphosphorous lipids such as sulfolipids under phosphate limitation (Van Mooy et al. 2009).

Still, we know specific biological functions only for comparatively few lipids. The in-depth functional studies that have been performed by molecular and cellular biologists on certain lipid types [e.g., photosynthetic pigments such as carotenoids (Frigaard et al. 2004, Vogl & Bryant 2012); retinal, the prosthetic group of the light-absorbing protein rhodopsin (Kiser et al. 2012); and signaling molecules such as steroid hormones (Miller & Auchus 2011)] provide examples of the level of commitment required to gain deep knowledge about lipid function. Geobiologists interested in interpreting the molecular fossil record rigorously should be inspired by such studies and must consider the varied physiological roles that specific molecular fossils may once have played in living cells.

4. DISTRIBUTION PATTERNS OF HOPANOIDS AND HOPANES

The rock record hosts a vast amount of ancient carbon in the form of molecular fossils, most of which are so-called orphan biomarkers [i.e., compounds for which structures but not source organisms are known, such as cheilanthanes and 17,21-secohopanes (Aquino Neto et al. 1982, 1983)]. Collectively, they record a detailed history of ancient life, yet at present we can decipher only a few vague clues. A substantive interpretation of any molecular fossil requires three things: (*a*) knowledge of its occurrence pattern (where and when was it deposited?), (*b*) understanding of the biology of its molecular progenitor (by which organisms and under what influences was it made, and how?), and (*c*) awareness of how it may have been altered during its fossilization.

By way of example, we focus on the occurrence pattern of a class of triterpenoid lipids known as hopanoids. Their diagenetic products, hopanes, are among the longest-lived molecular fossils and are ubiquitous in modern and ancient systems (**Figure 3**). We briefly review how hopane abundance patterns correlate with other geochemical proxies, and how these correlations have been interpreted. We must remember that correlation is not causation, and that these interpretations represent hypotheses that require testing; we share an example of how we tested one such hypothesis in the following section. For the remainder of this review, hopanes and hopanoids serve as an example from which we draw general principles to guide biomarker interpretation.

4.1. Sources and General Distribution

It has been estimated that the mass of hopanes in sedimentary rocks and oil reservoirs is similar to the mass of organic molecules in all living organisms combined (Ourisson & Albrecht 1992). For many years, hopanes belonged to the enigmatic orphan biomarker family. This changed with the discovery of the parent compounds in *Acetobacter xylinum*, which led to the proposal that bacterial hopanoids are functional analogs to eukaryotic sterols, such as cholesterol, based on their structural similarities (**Figure 3**, structure III), which assist in rigidifying membranes (Förster et al. 1973, Rohmer & Ourisson 1976, Rohmer et al. 1979). Although hopanoids are present in approximately 10% of bacterial species and have an irregular phylogenetic distribution (Pearson et al. 2007, Frickey & Kannenberg 2009), they are widespread in diverse contemporary environments (Talbot et al. 2003, Farrimond et al. 2004, Pearson & Rusch 2009, Xu et al. 2009,

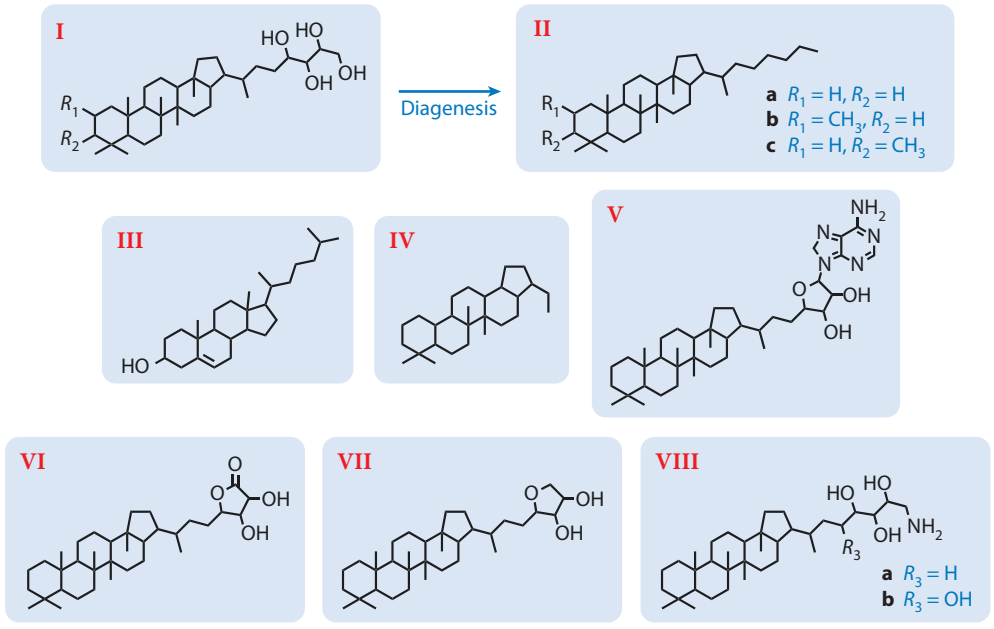


Figure 3

Structures of select hopanoids found in extant organisms, modern environments, and ancient sedimentary depositions. During diagenesis, hopanoids become fossilized hopanes and lose many of their differentiating structures, except C-2 and C-3 methylations. The structures illustrated in each panel are as follows (note R_1 and R_2 groups are the same for panels I and II): Ia, bacteriohopanetetrol; Ib, 2-methylbacteriohopanetetrol; Ic, 3-methylbacteriohopanetetrol; IIa, C_{35} hopane (i.e., pentakishomohopane); IIb, C_{35} 2-methylhopane; IIc, C_{35} 3-methylhopane; III, cholesterol; IV, 25,28,30-trisnorhopane; V, adenosylhopane; VI, hopaneribonolactone; VII, 32,35-anhydrobacteriohopanetetrol; VIIIa, aminobacteriohopanetriol; VIIIb, aminobacteriohopanetetrol.

Sáenz et al. 2011a, Zhu et al. 2011) and throughout geologic time (Peters et al. 2005). Because certain hopanoid types exhibit specific distribution patterns, they bear potential to provide insight into past ecological and environmental conditions (Summons et al. 1999).

During diagenesis, hopanoid parent molecules lose many of their differentiating features, such as hydroxyl and amine groups, leaving behind only their hydrocarbon skeletons. In contrast, modifications that can be preserved include methylations to the hydrocarbon backbone at C-2 or C-3 (**Figure 3**). Methylated hopanoids are important biomarkers because methylation allows different hopanes to be distinguished from each other in the fossil record. 2-Methylhopanes (e.g., C_{35} 2-methylhopane) (**Figure 3**, structure IIb), the preserved form of 2-methylhopanoids (e.g., 2-methylbacteriohopanetetrol) (**Figure 3**, structure Ib), are among the oldest syngenetic biomarkers, dating to 1.64 Ga in the Barney Creek Formation, an anoxic sulfidic permanently stratified marine depositional setting (Brocks et al. 2005). Furthermore, the 2-methylhopane index, the ratio of 2-methylhopanes to methylated and unmethylated hopanes, shows intriguing patterns of change over time. Relatively high 2-methylhopane indices are found throughout the Proterozoic (Summons et al. 1999), and the highest measurements co-occur with Oceanic Anoxic Events (OAEs), including at the Permo-Triassic boundary (Xie et al. 2005) and during Cretaceous OAEs (Kuypers et al. 2004) (**Figure 1*b***). OAEs are characterized by depletion of oxygen in the oceans but are also associated with global disruptions in nutrient cycling and mass extinctions. Isotope data suggest that the nitrogen cycle was disrupted during OAEs, enhancing

the need for fixed nitrogen and causing major changes to the nitrogen cycle (Kuypers et al. 2004, Knoll et al. 2007, Higgins et al. 2012, Monteiro et al. 2012).

What might these distribution patterns reveal about ancient life? In general, hopanoids in modern sediments and hopanes in the geologic record have been used as indicators for bacterial activity or bacterial input to sediments. Numerous minor structural variations within the molecular class can be linked to more particular functional, taxonomic, or environmental interpretations. For example, certain hopanoids have been suggested to be diagnostic tools for sedimentary diagenesis; tracers of soil, terrestrial, or estuarine carbon; and indicators of methanotrophic activity. We summarize these ideas below but reserve interpretation of 2-methylhopanes for the following section.

4.2. Indicators of Sediment Diagenesis

Although diagenesis can preserve critical aspects of the original skeletal structure (e.g., 2-methylhopane), other parts of the molecule are often altered in ways that are both predictable and diagnostic. The ratio of total hopanes to total steranes has been used to interpret changes in the fraction of bacterial versus algal input to sediment and petroleum (Moldowan et al. 1985, Sinninghe Damsté & Schouten 1997, Andrusevich et al. 2000). An increased abundance of norhopanes (e.g., 25,28,30-trisnorhopane) (**Figure 3**, structure IV) indicates a direct hydrocarbon contribution to sediments, because these compounds are products of biodegradation and are not produced during thermal maturation (Noble et al. 1985). Hopane chain length proxies such as the C_{35} homohopane index, defined as $[C_{35}/(\Sigma C_{31-35})]$, record the extent of degradation of hopanoids. A higher index, denoting less degradation, is believed to reflect greater anoxia (Peters & Moldowan 1991). Together these various proxies are used as tools to interpret the environmental conditions under which ancient sediments were deposited. For example, a high homohopane index accompanies other indicators of a biotic, oxygen-poor crisis associated with the Permo-Triassic extinction (Hays et al. 2012).

4.3. Tracers of Terrestrial Environments

In recent sediments, intact hopanoids may be useful tracers for carbon source inputs. Although few hopanoids are presently thought to be environmentally or metabolically specific, adenosylhopane (**Figure 3**, structure V) may be an exception. It is abundant in terrestrial and lacustrine environments, whereas it is usually below detection in sediments with negligible terrigenous input (Talbot & Farrimond 2007, Cooke et al. 2008, Pearson et al. 2009, Xu et al. 2009). However, because adenosylhopane is the first biosynthetic intermediate in the formation of all hopanoids with extended side chains (Bradley et al. 2010, Welander et al. 2012), the accumulation of adenosylhopane cannot be taxonomically diagnostic. It remains unknown why this apparently arrested synthesis would be more prevalent in soil-dwelling microbes, but it nonetheless appears to be a good tracer for terrestrial input. Similarly, the side-chain structures of hopaneribonolactone (**Figure 3**, structure VI) and 32,35-anhydrobacteriohopanetetrol (**Figure 3**, structure VII) have been found in oxidizing and reducing environments, respectively (Bednarczyk et al. 2005, Talbot et al. 2005, Sáenz et al. 2011b). Bradley et al. (2010) suggested that both compounds are generated abiotically through oxidative or reductive cleavage of an as-yet-unidentified precursor. The ratio of hopaneribonolactone to 32,35-anhydrobacteriohopanetetrol might provide a useful environmental or physiological signal, albeit independent of any relationship to taxonomy.

4.4. Markers of Methanotrophy

Among hopanoids that may have taxonomic associations, the best examples may be those that contain amine groups. Aminobacteriohopanetriol (**Figure 3**, structure VIIIa) has been found only in

methanotrophs and *Desulfovibrio* spp. (Blumenberg et al. 2006), and Type I methanotrophs known to contain aminobacteriohopanetetrol (**Figure 3**, structure VIIIb) are observed in areas where aerobic methane oxidation is an important component of the carbon cycle (Talbot et al. 2003, Zhu et al. 2010, van Winden et al. 2012). However, the best confirmation of a methanotrophic signal is the simultaneous evidence of depleted $\delta^{13}C$ values that can be found in sedimentary hopanoids, many of them being the potential diagenetic products of these functionalized precursors. Such examples are primarily associated with lacustrine systems (Freeman et al. 1990) or marine methane seeps (Elvert & Niemann 2008). Interestingly, this signal also occurs in anoxic sediments associated with anaerobic oxidation of methane (Thiel et al. 2003). Differentiation between aerobic and anaerobic oxidation of methane therefore has relied on the presence of 3-methylhopanoids, markers for aerobic and/or acetic acid bacteria (Zundel & Rohmer 1985) (**Figure 3**, structures Ic and IIc), although these compounds perhaps have a wider taxonomic and ecological distribution (Welander & Summons 2012).

5. CASE STUDY: GAINING INSIGHT INTO 2-METHYLHOPANOIDS

To determine whether causal linkages underpin correlations between molecular (fossil) distribution patterns and particular (paleo)environmental events, an experimental approach is needed. This is where molecular and cellular biologists can help advance biomarker interpretation. Though experiments performed with modern organisms cannot unambiguously demonstrate causality about ancient processes, they are the best we can do short of inventing a time machine. In this section, we provide an example of how one might approach validating functional interpretations of a molecular fossil; specifically, we describe the sequence of questions we asked and tests we performed to gain insight into the biological function of 2-methylhopanoids (**Figure 4**). Though more remains to be done before we can tell a complete story, significant progress has been made during the past decade in how to interpret 2-methylhopane distribution patterns, illustrating how molecular and cellular biological approaches can be applied to such problems.

5.1. Questioning the Interpretation of 2-Methylhopanes as Proxies for Oxygenic Photosynthesis

Early studies of bacterial culture collections found evidence for hopanoid production primarily in bacteria growing aerobically (Rohmer et al. 1984), suggesting that geologic hopanes were good general indicators for oxygenated environments. Since then, hopanoids have been reported in a wide range of bacterial species growing anaerobically and have been found in diverse environments, demonstrating that there is not a requisite connection between hopanoids and oxic ecosystems or habitats (Fischer et al. 2005, Härtner et al. 2005, Blumenberg et al. 2006, Eickhoff et al. 2013). Moreover, the distinctive 2-methylhopane skeleton was initially determined to be common in freshwater and mat-dwelling cyanobacteria (Summons et al. 1999) (**Figure 3**, structure IIb). Thus, when 2-methylhopanes were detected in 2.7 Ga strata, they were interpreted as evidence for the ancient evolution of oxygenic photosynthesis (Brocks et al. 1999). Though these deposits are now known to be nonsyngenetic contaminants (French et al. 2015), we questioned this interpretation for independent reasons. First, cyanobacteria are not the only organisms that can make 2-methylhopanoids, and not all cyanobacteria produce them (Summons et al. 1999, Sáenz et al. 2012b), indicating there is neither a unique nor a consistent correlation between 2-methylhopanoid production and cyanobacteria. Second, no evidence existed that directly linked 2-methylhopanoid production to the process of oxygenic photosynthesis. Because cyanobacteria can ferment to survive (Stal 1991, Margheri & Allotta 1993) and generate survival structures

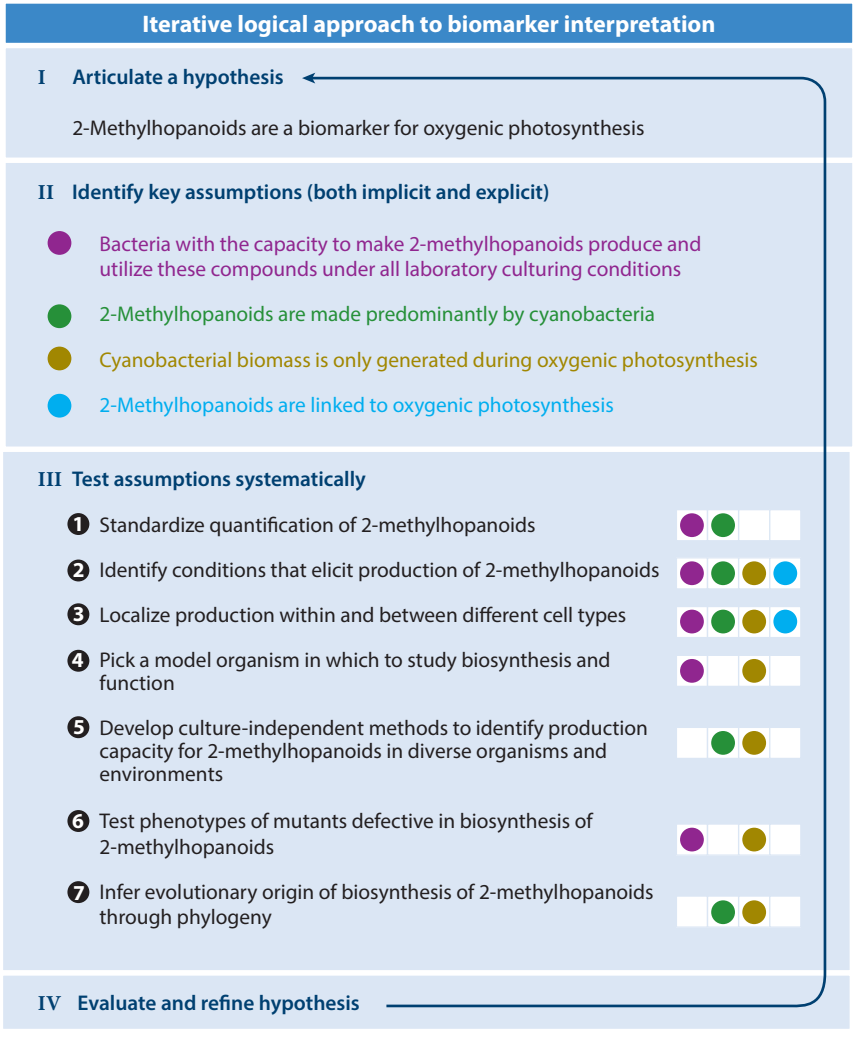


Figure 4

Iterative logical approach to biomarker interpretation. Four basic steps are required to test biomarker interpretations. We provide examples of each step related to the interpretation of 2-methylhopanes. Color is used to indicate approaches used to test specific assumptions.

in the absence of light (Doughty et al. 2009), we could not assume that 2-methylhopanoids are made only when cyanobacteria are generating oxygen. We therefore set out to test the hypothesis that 2-methylhopanes are proxies for oxygenic photosynthesis using bacteria in which we could determine the physiological functions of 2-methylhopanoids.

5.2. Laying the Groundwork: Analysis and Characterization

Before testing hypotheses about potential functions for a biomarker, it is vital to carefully characterize its occurrence pattern among and within different cell types. This requires not only precise

and accurate quantification methods but also awareness of where to look, at both the cellular and subcellular levels. We structure the following subsections around a set of generic questions we used when studying 2-methylhopanoids: Which lipids are present and in what abundance? Who produces the lipid of interest? Where do particular lipids reside within different cell types, and with what do they interact?

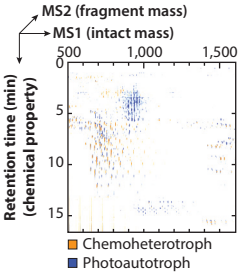
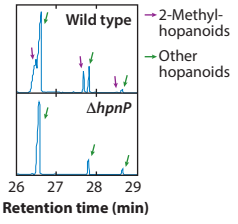
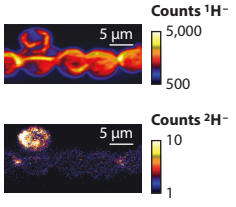
5.2.1. Which lipids are present and in what abundance? Significant progress has been made in developing protocols for assessing the syngeneity of biomarkers in ancient rocks (Brocks 2011, French et al. 2015), enabling more accurate measurements of biomarker distribution patterns. Similarly, analytical accuracy is needed when studying the lipids that ultimately become biomarkers (**Table 1**). The ability to quantitatively measure structural variants enables a range of follow-up studies, from detecting their production by different organisms or within different membranes of a given organism to understanding what stimulates their production to gaining insight into their biophysical or biochemical functions. In addition, absolute quantification of biomarkers permits meaningful comparisons between studies, allowing more accurate hopanoid pattern detection in both laboratory and environmental samples. A recent step forward in quantitative hopanoid analysis has come from the synthesis of isotopically labeled internal hopanoid standards (Wu et al. 2015).

5.2.2. Who produces the lipid of interest? With accurate analytical methods in place, one can determine whether a specific lipid is diagnostic for a particular taxon or a group of phylogenetically related organisms or is made by members of diverse clades. Even when a lipid is not specific to a group of organisms, knowing which species can produce it can influence thinking about its function and can help identify the best model organism in which to study it (see sidebar). With the unexpected discovery that an anoxygenic phototroph, *Rhodopseudomonas palustris* TIE-1, could produce 2-methylhopanoids in equal amounts as cyanobacteria [earlier reports had either not detected 2-methylhopanoids in related species or detected them only at very low levels (Rohmer et al. 1984, Zundel & Rohmer 1985, Knani et al. 1994, Summons et al. 1999)] it became clear that accurate identification of 2-methylhopanoid-producing organisms was sensitive to the culturing condition (Rashby et al. 2007). Though biologically interesting, such condition-dependent biomarker production is problematic from the perspective of identifying source organisms, as biomarkers might easily be missed if a nonstimulatory culturing condition were used. Accordingly, we sought a culture-independent method to predict the capacity for 2-methylhopanoid production.

THE VALUE AND LIMITS OF MODEL ORGANISMS

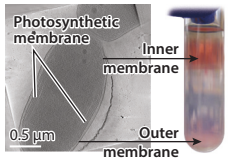
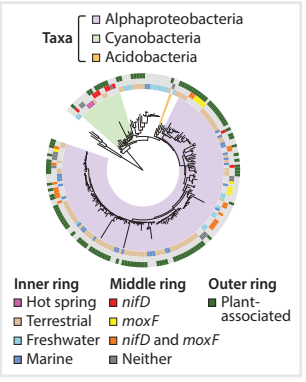
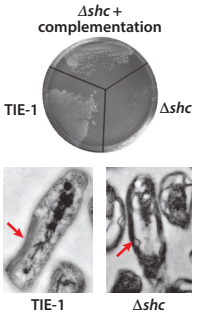
The selection of a good model organism can greatly facilitate progress in identifying the biosynthetic pathway and biological functions of a lipid biomarker. Because biosynthetic pathways are often conserved, research is expedited when an organism is used that is easy to culture and amenable to genetic manipulation. Ideally, this organism should produce the lipid in appreciable quantities under certain conditions but not require it in others, have a sequenced genome, and occupy an ecological niche that is readily simulated in the laboratory. Conditional lipid production facilitates the generation of mutants that are unable to produce the biomarker. Mutant strains can then be used to identify conditions under which its production provides a fitness advantage, provided other molecules are unable to functionally substitute for the lipid of interest. Though functional studies in model organisms provide an excellent starting point for biomarker interpretation, it is important to determine whether functions identified in any particular organism also extend to different strains before making broader conclusions.

Table 1 Common approaches used to study lipid functions

Question	Technique	Visual example
What? How much?	**Liquid chromatography mass spectrometry (LC-MS).** LC-MS is particularly useful in analyzing diverse lipid types. It combines the large separation capacity of ultraperformance liquid chromatography (UPLC) with mass spectrometry (MS). Often time-of-flight (TOF), ion trap (e.g., Orbitrap) and/or quadrupole MS are used. Qualitative information on a large number of metabolites can be collected in a single run. For quantitative analysis of particular lipid types, sample complexity can be reduced before analysis (e.g., by solid phase extraction), deuterated internal standards can be added—if available—and the ionization method can be varied [electrospray ionization (ESI), chemical ionization (CI), photoionization (PI), among others, all performed in either the positive- or negative-ion mode]. Retention time and fragmentation scans carry structure-specific information. The ionization condition can interfere with authentic sample representation due to reactivity, ion suppression, and other effects. For biological samples, **http://www.lipidmaps.org** provides excellent protocols.	Overlay of LC-MS total ion chromatograms of lipid extracts of *Rhodopseudomonas palustris* grown in two different conditions (Neubauer et al. 2015). 
	Gas chromatography mass spectrometry (GC-MS). GC-MS is widely used for the analysis of geological biomarkers and other volatile compounds. For nonvolatile samples such as polar lipids and fatty acids, chemical derivatizations are available to modify their polar functional groups and lower boiling points for analysis. GC-MS has high sensitivity, wide dynamic range, and relatively low cost. The most common MS for GC is a single-quadrupole mass analyzer with an electron ionization source. For most accurate quantification, both external and internal standards are required to account for variations between runs and the fluctuation in the sensitivity of the MS.	GC-MS total ion chromatograms of acetylated total lipid extracts of *R. palustris* cultures (Welander et al. 2010). 
Where?	**Mass spectrometry imaging (MSI).** Mass spectrometry–based imaging techniques can resolve metabolites or isotopes within tissues and cells, where understanding spatial variation is important. Approaches based on matrix assisted laser desorption (MALDI) and nanoscale secondary ion MS (nanoSIMS) require dehydration of the sample. Ambient ionization techniques, for example based on desorption ionization (DESI), can be used on live samples such as microbial colonies, but at much lower resolution (>10 μm).	Hopanoid localization in Nostoc with NanoSIMS (Doughty et al. 2014). 

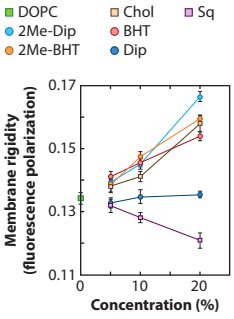
(*Continued*)

Table 1 (*Continued*)

Question	Technique	Visual example
	Membrane fractionation. Despite their small size, microbial cells are highly organized. Most biomolecules are not homogeneously distributed throughout the cell but are instead concentrated in specific locations. A classic biochemical approach is to fractionate membranes by their types, such as separating the inner and outer membranes of gram-negative bacteria. Membrane fractionation typically involves cell disruption by sonication, shearing forces, osmotic pressure, or chemical lysis, followed by ultracentrifugation of the sample through a high-density solution or gradient (e.g., sucrose, Percoll, PEG/Dextran) to separate distinct membrane types. Acceptable homogeneity of these preparations is often tested by enzymatic assays that track proteins residing in only one membrane type, or by quantifying diagnostic proteins using antibodies (Western blot).	Membranes in *R. palustris* and biochemical fractionation (Neubauer et al. 2015, Wu et al. 2015). 
What does it do?	**Phylogenetic analysis.** Genomic and metagenomic DNA sequences contain information about phylogenetic relationships. Once a gene (and hence its DNA sequence) that is involved in biomarker biosynthesis or utilization is known, it is possible to search databases to learn which organisms possess related genes, which may have common functions. If sufficient sequences similar to the originally identified gene can be identified, primers (i.e., short oligonucleotides) can be designed to pair with sequences shared by all members of the gene family, thereby enabling the amplification of related genes via the polymerase chain reaction (PCR) from laboratory or environmental samples. Clone libraries created from these new sequences can assist in the identification of more extensive sequencing of distant gene variants, which aids phylogenetic analysis. Different methods, ranging from microfluidic digital PCR (Ottesen et al. 2006) to single-cell genome sequencing (Kalisky et al. 2011), can help determine which strains encodes these novel variants. Resources: **http://www.microbesonline.org**, **http://img.jgi.doe.gov**.	HpnP diversity from metagenomes and its correlation with plant-microbe interactions (Ricci et al. 2014). 
	Gene deletion. Gene deletion permits testing of whether a gene is required for a particular process in vivo, such as the biosynthesis of a specific lipid. Some microorganisms are more amenable to genetic analysis than others (see sidebar and Related Resources). To determine whether the effects in vivo are indeed caused by the loss of a particular gene and not by some random mutation, complementation [expression of a clean copy of the gene on a plasmid or (re)introduced on the chromosome] is attempted in the mutant strain. If the phenotype is restored by complementation, this proves the gene is necessary. Alternatively, one can test whether a gene is responsible for a function by overexpressing it in the native organism or expressing it in a different organism that does not have the gene; if the function is enhanced or gained, respectively, this indicates causality. Additional insights can be gleaned from studying mutations that enable mutants to adapt to a gene loss (e.g., suppressor mutations). A wide array of approaches to screen for new phenotypes or select for a defined property can be applied. Other in vivo assays based on using fluorescently labeled proteins [e.g., green fluorescent protein (GFP)] or dyes that provide information about membrane charge or the existence of microdomains are just a few examples of the many ways one can gain complementary insights into gene function.	A hopanoid mutant (Δ*shc*) has membrane defects and is susceptible to bile salts compared to the wild-type strain (TIE-1) (Welander et al. 2009). 

(*Continued*)

Table 1 (*Continued*)

Question	Technique	Visual example
	In vitro assays. A goal of many biochemical studies is to identify the minimum set of pure components to reconstitute a process in a test tube. In the realm of lipids, the effects of a particular lipid type, such as its impact on membrane rigidity, permeability, curvature, melting temperature, etc., can be measured in vitro with reconstituted membranes or vesicles. The most widely used techniques utilize vesicles, which form when solutions of lipids are pushed through fine pores. For example, with the aid of a fluorescent reporter dye, membrane fluidity can be estimated by fluorescence polarization. Similarly, the rigidity of whole cells can be determined and compared between different (mutant) strains.	Membrane rigidity measurements using model lipids (Wu et al. 2015). 

Genes involved in the biosynthesis of biomolecules are often conserved. We therefore sequenced the genome of *R. palustris* and set out to identify the gene responsible for methylation at C-2. Bacteria often group genes of the same pathway together on the chromosome, a phenomenon referred to as gene synteny. The enzyme squalene hopene cyclase, encoded by the gene *shc* (Wendt et al. 1997), was known to catalyze the first step in hopanoid biosynthesis (Seckler & Poralla 1986), so we identified the genomic region containing *shc* using homology searching (i.e., the use of sequence similarity to identify hopanoid biosynthetic capacity). In close proximity to *shc* were two genes whose products were predicted to belong to a class of enzymes capable of performing methylation reactions in the absence of oxygen. Deletion of one of these loci resulted in the loss of 2-methylhopanoid production; we named this locus *hpnP*. Provision of a fresh copy of the *hpnP* gene to the deletion strain restored 2-methylhopanoid production (Welander et al. 2010). Bacterial geneticists call such an experiment complementation analysis, and it is the gold standard for demonstrating causality because it demonstrates that the missing gene, rather than a mutation elsewhere on the chromosome, is responsible for the phenotype of interest (**Table 1**).

Identifying the gene or genes that are responsible for the biosynthesis of a particular lipid opens up the possibility of identifying potential source organisms by homolog searching. Of course, homology does not prove that two enzymes have the same function, and search tools are limited by what is in the database. Nevertheless, HpnP, the enzyme encoded by *hpnP*, has a well-conserved amino acid sequence that is readily distinguishable from that of related enzymes, and all 2-methylhopanoid producers so far described have *hpnP* (i.e., no organism without *hpnP* has yet been found to methylate hopanoids). Accordingly, *hpnP* presence is a robust, culture-independent predictor of an organism's ability to make 2-methylhopanoids. HpnP homology searches in sequenced bacterial genomes revealed that a minority of sequenced cyanobacteria make 2-methylhopanoids, whereas a large number of alphaproteobacteria (concentrated in the Rhizobiales clade) and an acidobacterium have this capacity (Welander et al. 2010, Ricci et al. 2015). By amplifying *hpnP* from environmental DNA and searching for *hpnP* in various metagenomic databases, we found further evidence that organisms other than cyanobacteria are the dominant producers of 2-methylhopanoids in diverse environments (Ricci et al. 2014) (**Table 1**). Notably, the one location we surveyed in which culture-independent techniques indicated that

cyanobacteria significantly contribute to 2-methylhopanoid production in hot spring microbial mats was Yellowstone, historically an important site of culture-based efforts to identify 2-methylhopanoid producers (e.g., Zhang et al. 2007, Talbot et al. 2008).

Today it is possible to scan many environments quickly and easily by searching for the presence of marker genes among publicly available (meta)genomic data sets. In doing this for *hpnP*, we found that metagenomes from plant-containing environments harbor significantly more *hpnP* sequences than do those from other environments. We also noticed that *hpnP* is enriched among plant-associated bacteria, and strongly correlates with the presence of genes known to be required for metabolisms that underpin plant-microbe interactions (e.g., *nifD*, involved in nitrogen fixation, and *moxF*, involved in methanol utilization). In addition, we observed diverse *hpnP* types in closely packed microbial communities from other environments, including stromatolites, hot springs, and hypersaline microbial mats. The common features of these niches indicate that 2-methylhopanoids are enriched in sessile microbial communities whose habitats are low in oxygen and fixed nitrogen and have high osmolarity.

5.2.3. Where do particular lipids reside, and with what do they interact? Refining assumptions about who makes a particular lipid is an important first step toward identifying a good model system in which to study its function. Similarly, awareness that organisms themselves can exist as different cell types, that within any given cell there can be multiple membranes, and that membranes are heterogeneous mixtures of lipids and proteins provides an additional framework for correlative analyses (**Table 1**). Again, we underscore that correlation is not causation, yet collecting correlative data at multiple scales can suggest testable hypotheses about function. For example, some 2-methylhopanoid-producing cyanobacteria make filaments with distinct cell types. *Nostoc punctiforme* forms vegetative cells, heterocysts (dedicated to nitrogen fixation), and akinetes (spore-like survival structures). Qualitatively, 2-methylhopanoids are present in the greatest abundance in the outer membranes of akinetes (Doughty et al. 2009), suggesting that their function in these organisms may have more to do with enhancing stress resistance than with metabolism. However, as described above, lipids can have multiple functions and may indirectly affect a variety of cellular processes. It can therefore be helpful to catalog positive or negative correlations at the molecular level between the lipid of interest and other lipids or proteins, as linked abundance patterns may hint at potential functions. For instance, 2-methylhopanoids might improve the efficiency of a membrane transporter via direct interactions. Or 2-methylhopanoids might functionally substitute for a different type of lipid, resulting in the lowering of its abundance when 2-methylhopanoids are present. Alternatively, 2-methylhopanoid biosynthesis might shift metabolite pools so that other lipids are produced at lower levels. There are many possibilities, and determining whether such putative molecular interactions are relevant, and if so, how they work, requires functional analyses that move beyond correlations.

5.3. How to Determine the Biological Function(s) of a Lipid

Deducing what a particular lipid type does in a cell can be very challenging, especially if one is attempting to understand a subtle chemical modification, such as methylation at C-2. Yet one can stack the deck in one's favor by starting with the simplest model organism that is amenable to genetic analysis. Genetic tractability is critical because genetics uniquely permits assessment of in vivo function. As much as biochemistry and biophysics can contribute to a deep mechanistic understanding of biological processes, without the ability to generate mutants, it is difficult to know whether in vitro experiments are physiologically relevant. Here we summarize the approaches we have taken to gain insight into the biological functions of 2-methylhopanoids.

5.3.1. Pursuit of a phenotype in *Rhodopseudomonas palustris* TIE-1. Because 2-methylhopanoids are produced by diverse bacteria, determining their cellular role is likely to provide greater insight into 2-methylhopane interpretation than merely refining their taxonomic associations. Although the function of a biomarker may change over time, molecular functions are generally more conserved than gene presence or absence, especially given the prevalence of horizontal gene transfer (Kunin et al. 2005). One way to decipher a lipid's function is to delete the gene responsible for its production and then study the resulting phenotype (e.g., properties) of the mutant strain. Sometimes phenotypes are easy to identify, particularly if the functions encoded by the gene are unique, but often defects are subtle and may only be visible under particular conditions or when multiple mutations are made to circumvent redundancy.

The mutants generated during our efforts to define the hopanoid biosynthetic pathway in *R. palustris* enabled a variety of other functional studies. These included (*a*) identification of a role for hopanoids in stress tolerance using the Δ*shc* strain (the Δ symbol is used by bacterial geneticists to denote deletion of a gene, i.e., this strain lacks *shc* and thus cannot convert squalene to diploptene, the first step in hopanoid biosynthesis; Δ*shc* does not produce any hopanoids) (Welander et al. 2009), (*b*) demonstration that HpnN transports hopanoids from the inner to the outer membrane by quantifying hopanoid abundance in membranes of the Δ*hpnN* strain compared to the wild type (Doughty et al. 2011), (*c*) determination that structurally different hopanoids differentially contribute to outer and inner membrane permeability by comparing bile salt resistance in the presence or absence of EDTA by the Δ*shc*, Δ*hpnP*, Δ*hpnH* (unable to make C_{35} hopanoids), and Δ*hpnO* (unable to make hopanoid aminotriol) strains relative to the wild type (Welander et al. 2012), and (*d*) recognition that 2-methylhopanoid biosynthesis is triggered by the general stress response (GSR) pathway through measurement of *hpnP* transcripts in various mutant backgrounds lacking the GSR machinery relative to the wild type (Kulkarni et al. 2013).

Though these studies with *R. palustris* advanced our understanding of hopanoid biology, despite our best efforts, we could never find a condition where the Δ*hpnP* strain showed a strong phenotype. In other words, the absence of 2-methylhopanoids did not appear to impact *R. palustris* significantly. There were two possibilities: (*a*) Other lipids might functionally substitute for 2-methylhopanoids in the Δ*hpnP* strain, thus obscuring our ability to see a phenotype, and/or (*b*) we hadn't yet found the optimal experimental conditions (appropriate growth regime) in which to detect a phenotype. While acknowledging the null hypothesis (2-methylhopanoids might be functionally insignificant), we proceeded to test these options.

5.3.2. Studies of biophysical effects and regulation. After improving our ability to detect hopanoids in the background of the greater lipidome, we realized that significant lipid remodeling can occur in hopanoid mutant strains (both Δ*shc* and Δ*hpnP*) under certain growth conditions (Neubauer et al. 2015). Moreover, the total amount of 2-methylhopanoids made by *R. palustris*, even under optimal growth conditions, constitutes a relatively small fraction (1–3 mol%) of the total lipid extract (Wu et al. 2015). These observations suggested that detecting a biological function for 2-methylhopanoids was likely to be hard, if not impossible, through phenotypic analysis of single mutant strains. Accordingly, we took an in vitro approach to test the effects of hopanoid 2-methylation upon membrane biophysical properties (Wu et al. 2015). Previously, in vitro experiments using model lipids indicated that bacteriohopanetetrol cyclitol ether can condense membranes at high temperatures but fluidize membranes at low temperatures (Poralla et al. 1980). Similarly, bacteriohopanetetrol and bacteriohopanemonol can condense model membranes (Kannenberg et al. 1983, Ourisson & Rohmer 1992), and diplopterol can form liquid ordered microdomains (Sáenz et al. 2012a) and preferentially interact with saturated lipids such as Lipid A (Sáenz et al. 2015). By performing reconstitution experiments with native

membranes of hopanoid-mutant strains, we were able to show that methylation at C-2 specifically increases the rigidifying effect of hopanoids, possibly due to improved packing within the phospholipid milieu (Wu et al. 2015) (**Table 1**).

Further insight into the function of a biomarker can come from understanding how the genes involved in its biosynthesis are regulated. A starting point for examining regulation is to determine when a gene of interest is transcribed. By examining sequences upstream of *hpnP*, we identified a conserved motif that was known to bind a transcription factor associated with the GSR pathway in related organisms; genetic analysis proved that the GSR pathway is required for *hpnP* expression in *R. palustris* (Kulkarni et al. 2013). However, we noticed that the amount of *hpnP* expression in our experiments did not directly match the amount of 2-methylhopanoids produced. This suggested that 2-methylhopanoid biosynthesis was regulated by factors we had not yet tested. Intriguingly, the GSR machinery is found in other 2-methylhopanoid-producing alphaproteobacteria, such as *Bradyrhizobium diazoefficiens* and *Methylobacterium extorquens*, in which it is important for establishing and maintaining their symbioses with plants. This connection seemed significant in light of our earlier finding that *hpnP* presence is significantly correlated with organisms, metabolisms, and environments that support plant-microbe interactions (Ricci et al. 2014, 2015). To test whether hopanoids, and 2-methylhopanoids in particular, contribute to fitness under these conditions, we turned to *B. diazoefficiens* as a model organism.

5.3.3. Pursuit of a phenotype in *Bradyrhizobium diazoefficiens*. Unlike *R. palustris*, *B. diazoefficiens* has several defined plant partners, making it straightforward to study the impact of hopanoid production in plant-microbe symbiosis (i.e., a natural ecological context one can readily generate in the laboratory). Using this organism, for the first time, we were able to find a significant phenotype for the Δ*hpnP* mutant: defective growth in the free-living state under hypoxic conditions or low pH (Kulkarni et al. 2015). The fact that we found a phenotype in this organism but not in *R. palustris* may be related to the fact that 2-methylhopanoids constitute a large percentage of its total hopanoid pool under this condition (~77%), which cannot be easily replaced when absent. Intriguingly, we observed a greater range of defects for the Δ*hpnH* mutant: defective growth not only under hypoxia and low pH but also in the presence of high osmolarity, bile salts, oxidative stress, and antimicrobial peptides, as well as symbiotic defects with the host *Aeschynomene afraspera*, including lower rates of nitrogen fixation (Kulkarni et al. 2015). Intriguingly, the C_{35} hopanoid bacteriohopanetetrol is enriched in hypoxic portions of the marine water column (Kharbush et al. 2013), though whether it is generated in these regions remains to be shown.

5.3.4. Localizing (2-methyl)hopanoids within and between different cell types. To advance our understanding of how hopanoids contribute to fitness in relevant ecological contexts, an important next step is to identify where they localize and with what other biomolecules they interact. Many hopanoids localize to the outer membrane of gram-negative bacteria; they are conveyed from the inner to the outer membrane by the membrane transporter HpnN in *R. palustris*, and likely also in other hopanoid-producing bacteria (Doughty et al. 2011). In the outer membrane, some C_{35} hopanoids can be conjugated to Lipid A (Silipo et al. 2014), yet others appear to be free in the membrane. Molecular simulation of bacteriohopanetetrol and diplopterol in phospholipid bilayers suggests their orientation can be fine-tuned by their functional groups (Poger & Mark 2013), adding another layer of complexity. We note that the orientation of specific hopanoids within the different leaflets of either the inner or outer membrane has not yet been experimentally validated, nor have systematic studies been done to determine whether different hopanoids localize to particular microdomains or whether such localization is functionally

important. Such studies are technically challenging because it is difficult to visualize native lipids; however, we have learned a little about (2-methyl)hopanoid localization in distinct cell types of two different organisms.

Membrane fractionation experiments of different cell types of *N. punctiforme* showed that 2-methylhopanoids are most abundant in the outer membrane of akinetes (Doughty et al. 2009). Using an isotope-labeling approach and nanoscale secondary ion mass spectrometry (nanoSIMS) imaging, we confirmed (2-methyl)hopanoid enrichment in the akinete cell type and found that (2-methyl)hopanoids congregate between cells of *N. punctiforme* filaments (Doughty et al. 2014). At the time we performed these studies, we were unable to tell whether these patterns were functionally significant because we lacked an *N. punctiforme hpnP* mutant in which we could attempt metabolic complementation; this limitation has been rectified recently (Ricci 2015). Like *N. punctiforme*, *R. palustris* forms distinct cell types; *R. palustris* divides asymmetrically into sessile mother and motile swarmer cells. Intriguingly, we observed bacteriohopanetetrol to be present in low abundance in swarmer cells, but to increase in concentration as they develop into mother cells (Doughty et al. 2011); the biological significance of this abundance pattern is currently unclear.

It is tempting to speculate that the localization of 2-methylhopanoids to specific membranes might connect their cell biology to a larger physiological phenotype. For example, bacteriohopanetetrol and a phenylacetate-esterified hopanoid were found to be enriched in vesicle envelopes of a *Frankia* strain, a nitrogen-fixing organism, and it was proposed that hopanoids might act as an oxygen diffusion barrier for the oxygen-sensitive nitrogenase (Kleemann et al. 1994), yet later studies showed that hopanoids were not produced in response to nitrogen limitation (Nalin et al. 2000). Whether (2-methyl)hopanoids impact oxygen diffusion through the membrane, or whether they modulate the functioning of membrane proteins that support respiration or nitrogen fixation, remains to be determined. Identification and understanding of the molecular interactions mediated by hopanoids (e.g., effects on membrane permeability, functioning of membrane proteins) will help bridge functional biophysical studies and phenotypes of relevance to the rock record.

5.3.5. Assessing relevance: How reasonable is it to generalize? Implicit in the contention that the interpretation of a biomarker is best based upon the biological function of its lipid progenitor is the assumption that its function is conserved today and was the same in the past. Model organisms are just a starting point for making inferences, and it is critical to determine whether the functions for any given biomarker in any given organism are generalizable. Even if extant organisms utilize a lipid in different ways, a broader function based on its consensus role across many organisms may still be discernable. If, however, particular lipids have significantly different functions in varied organisms, nuanced biomarker interpretation will require taking other factors into consideration, such as environment or taxonomy, assuming such information is preserved in the rock record.

Though the biological role(s) for a given lipid must be tested in specific organisms one by one, its biosynthetic history is recorded in modern genomes. Phylogenetic analysis can reveal the taxon in which its biosynthesis originated and from what enzyme family the capacity to make it evolved, as well as whether the genes encoding its biosynthesis were subject to horizontal transfer, duplication, or gene loss events. The placement of a particular gene or protein on a phylogenetic tree can hint at its function, provided it clusters within a portion of the tree where functions have been experimentally demonstrated for a variety of related sequences. For example, phylogenetic analysis of the HpnP protein revealed that 2-methylhopanoids likely originated in the Alphaproteobacteria and that the gene required for their biosynthesis was transferred horizontally into the Cyanobacteria well after the invention of oxygenic photosynthesis (Ricci et al. 2015). Yet again we stress that our ability to infer accurate information about the evolutionary history

of a particular biomarker is only as good as the data sets we use that assign putative functions to sequence information.

5.4. Revisiting the Rock Record

Our research has revised the interpretation of the rock record with respect not only to 2-methylhopanes, but to hopanoid biomarkers in general. Understanding the general taxonomic distribution of hopanoid biosynthesis across microbial taxa and different environments (e.g., Pearson et al. 2007, Pearson & Rusch 2009, Ricci et al. 2014) points to greater diversity and abundance in terrestrial systems and microbial mats, and poorer abundance in the oxic, low-biomass-density open ocean. This general conclusion is consistent with our results specifically for 2-methylhopanoids and their role in mediating the types of environmental stresses commonly found in mats, soils, sediments, and other organic-rich systems.

Because we no longer can view 2-methylhopanoids as biomarkers of oxygenic photosynthesis, but rather as stress response indicators, a high 2-methylhopane content (and high hopanoid content in general?) likely indicates paleoenvironmental conditions where independent lines of evidence indicate that the capacity to respire under hypoxia and/or perform nitrogen fixation may have provided a selective advantage (Kuypers et al. 2004, Jenkyns 2010, Blumenberg et al. 2013, Kharbush et al. 2013). This suggests that during OAEs and other periods in Earth history in which organic-rich sediments accumulated 2-methylhopanoids, the compounds may reflect changes in sedimentary pH or redox conditions. Similarly, the hopanoid known as BHT isomer has been suggested to be a product of anaerobic ammonia oxidizers and thus local anoxia, although additional sources are also possible (Rush et al. 2014). Interesting in this regard is the lack of association of 2-methylhopanoids with blooms of *Azolla* ferns in the stratified, high-latitude Arctic Ocean of the Eocene, an environment believed to have experienced high rates of nitrogen fixation and which has bulk nitrogen isotope signals similar to those of OAEs (Speelman et al. 2009, Schoon et al. 2011). Two major unknowns for all of these cases remain the extents to which (*a*) deposited hopanoids could have been transported with terrestrial sediments, and/or are benthic (sedimentary) in origin, and (*b*) the sedimentary hopane record is a record of preservation rather than production. A sedimentary origin for hopanes is a reasonable suggestion, as indications are that hopanoid production by low-biomass-density planktonic ecosystems is low (Sáenz et al. 2011a), yet the molecular taphonomy of hopanoids (and many other biomarkers) is poorly constrained. Future studies examining the interplay of ocean redox, nitrogen fixation, and 2-methylhopanoid abundance may help tease out which environmental stressors are more influential on hopanoid distribution and preservation.

However, as we have emphasized throughout this review, correlation is not causation, and we must proceed cautiously when considering 2-methylhopanoids as biomarkers for any particular physiological process. Indeed, 2-methylhopanoids, and hopanoids more broadly, represent but one type of successful evolutionary adaptation to certain types of membrane stress, but cells have evolved other strategies to withstand them that are hopanoid independent. Not all bacteria that respire microaerobically or fix nitrogen make hopanoids, so something else has co-selected these attributes in particular cell types. Perhaps an evolutionarily serendipitous confluence of certain biosynthetic pathways modulates membrane fluidity and permeability well, leading to co-selection over time. Or perhaps a set of stresses that we have yet to define is found in certain environments but not others. For example, though the microenvironments of plant root nodules share some attributes (e.g., low pH, low O_2), hopanoids facilitate symbiosis only in certain plants; intriguingly, hopanoids confer a fitness advantage to bacterial symbionts of a plant that produces cysteine-rich antimicrobial peptides, yet appear to be expendable for symbionts of a plant that does

not produce these compounds (Kulkarni et al. 2015). Defining the specific combination of stressors that can explain an adaptive function for (2-methyl)hopanoids in some contexts but not others—i.e., a hopanoid/environmental stressor Venn diagram—is a priority for future research. The more we learn about what hopanoids interact with (or covary with) in the membrane, whether these relationships are conserved across multiple species, and whether specific chemical fingerprints

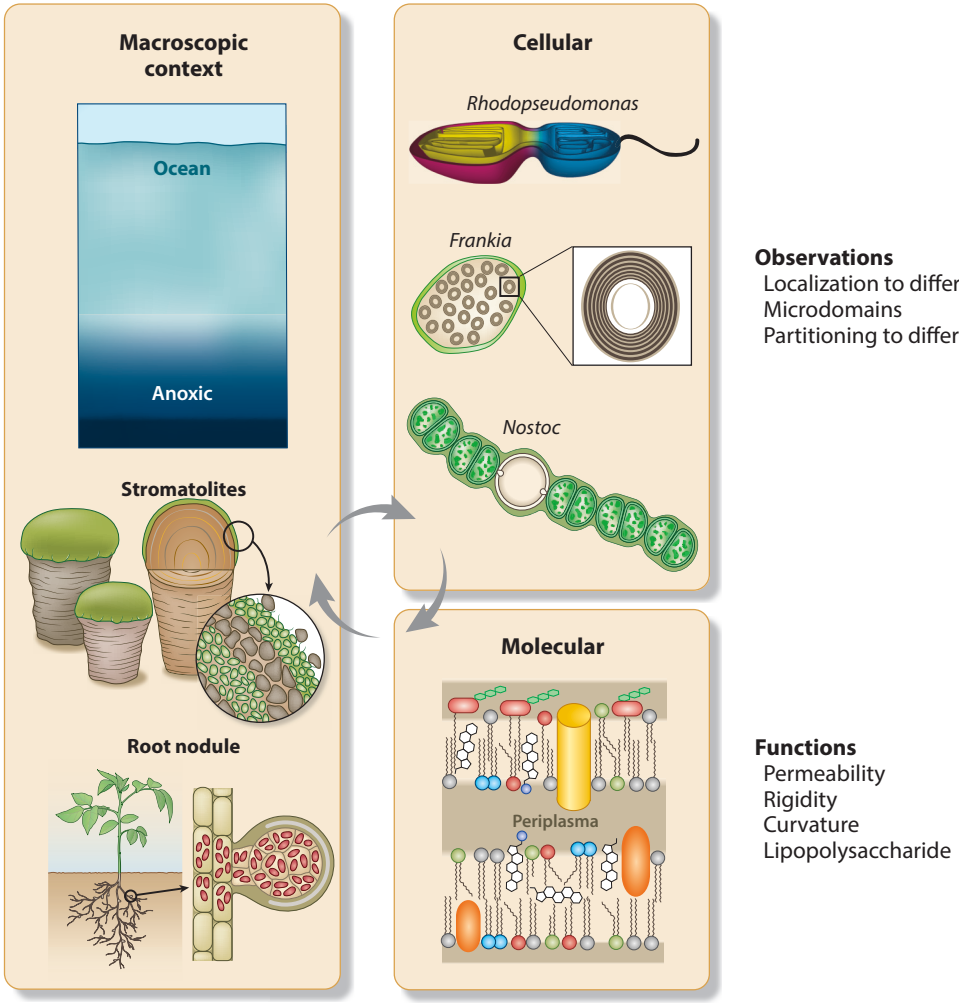


Figure 5

Bridging different scales in biomarker interpretation. The interpretation of biomarkers in the rock record requires studies on different scales. First, macroscopic data must be collected to identify modern and ancient environments where the biomarker is produced. In the case of (2-methyl)hopanoids, diverse contexts host (2-methyl)hopanoid-producing bacteria, from anoxic portions of the oceans to stromatolites to root nodules made by plant symbionts. Although they appear superficially different, at the microscale, the environments of these niches share many important attributes that may select for (2-methyl)hopanoid-producing organisms. To understand how cells respond to the environment by making (2-methyl)hopanoids, cellular studies are needed to reveal when they are produced and where they localize within and between different cell types. Ultimately, resolution at the molecular level is required to assess whether different hopanoids serve different functions, and whether these functions are conserved. Once such functions are understood, it may be possible to re-examine hopanoid occurrence at the macroscopic and cellular levels to articulate unifying hypotheses that help bridge modern experimental studies and 2-methylhopane distribution patterns in the rock record.

characterize the environments where (2-methyl)hopanoids are made by cells today, the better able we will be to interpret what their presence in modern or ancient environments implies. Though much remains to be done to bring the next chapter in the (2-methyl)hopanoid detective story to closure, the past decade has demonstrated how helpful taking a molecular and cellular biological approach can be to advancing a rigorous interpretation of an ancient biomarker.

6. CONCLUSION

The rock record is riddled with molecular fossils that have the potential to inform us about the history of life. In this review, we have attempted to explain the important role molecular and cellular biologists can play in helping to interpret this record. We have focused on 2-methylhopanes by way of example, yet the types of questions and logical approach illustrated by this story are generic and can be applied to other biomarkers. Ultimately, our ability to properly interpret the meaning of any given biomarker requires a detailed understanding of its function and distribution in modern microorganisms.

It is possible to bridge vast temporal and spatial scales when analyzing microbial biomarkers because the ecology of microorganisms is linked directly to their biochemistry via metabolic processes. Moreover, the microscale is a great equalizer: From a bacterial perspective, it matters less whether a cell is housed within a nodule, embedded within a stromatolite, or attached to particles of marine snow. What matters is the nature of its microenvironment, which can be highly conserved at the scale of the microbe despite different macroscopic contexts. It thus becomes possible to speculate on putative functions for ancient biomarkers by understanding the functions of their modern counterparts, provided these functions are conserved across different species and their niches share similar key properties (**Figure 5**).

Although the interpretation of molecular fossils is challenging and requires a commitment that spans many years, it is satisfying to be able to draw credible connections between modern molecular and cellular processes and biomarker abundance patterns in ancient rocks. We hope this review will inspire many such attempts by future geobiologists, so that one day the molecular fossil record may reveal more of its hidden stories about evolution in deep time.

SUMMARY POINTS

1. Significant progress in biomarker research has been made in the past decade due to improvements in biomarker sampling protocols and quantitative analyses and the expansion of sequence databases.
2. Correlation is not causation. The fact that today a lipid is made by a particular organism, or group of organisms, neither means that its biosynthetic ability evolved in members of this clade nor implies that its molecular fossils can be used to infer the presence of a specific taxon or metabolic process in paleoenvironments.
3. The lipid content of an organism is dynamic and can adjust its composition in response to different conditions. Thus, the absence of evidence for a particular lipid in a culture cannot be taken as evidence for the absence of an organism's ability to produce it.
4. Genetic analysis can help identify sequences that encode the enzymes responsible for lipid biosynthesis. Knowledge of these sequences can enable culture-independent surveys for lipid production capacity in modern environments, as well as phylogenetic analyses that can provide insight into their evolutionary histories.

5. The interpretation of ancient biomarkers is significantly strengthened by an understanding of the functions of their lipid antecedents, provided these functions are conserved.
6. The choice of a good model organism can accelerate progress to assess the physiological function of a particular lipid. Conclusions based on any given organism must be validated by comparative analyses across taxonomic divisions to determine the extent to which the conclusions are generalizable.
7. Tremendous opportunities exist for students trained in molecular and cellular biology to decipher the molecular fossil record.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This review is dedicated to Sky Rashby, who would have enjoyed knowing where his discovery led. D.K.N. is grateful to past and present members of the Newman laboratory for bringing their diverse talents to bear on a challenging problem, as well as to Alex Sessions and Roger Summons for an enjoyable long-term collaboration. We thank Nathan Dalleska and Stuart Conway for additional support with hopanoid analysis and synthesis, as well as Tanja Bosak, Andy Knoll, and Andreas Kappler for providing images used in **Figure 1**. Grants from the NASA Exobiology program (NNX12AD93G), National Science Foundation (1224158), and Howard Hughes Medical Institute (HHMI) enabled this research. D.K.N. is an HHMI Investigator. A.P. acknowledges support from the NASA Astrobiology Institute program and the Gordon and Betty Moore Foundation.

LITERATURE CITED

Allwood AC, Walter MR, Kamber BS, Marshall CP, Burch IW. 2006. Stromatolite reef from the early Archaean era of Australia. *Nature* 441:714–18

Andrusevich VE, Engel MH, Zumberge JE. 2000. Effects of paleolatitude on the stable carbon isotope composition of crude oils. *Geology* 28:847–50

Aquino Neto FR, Restle A, Connan J, Albrecht P, Ourisson G. 1982. Novel tricyclic terpanes (C_{19}, C_{20}) in sediments and petroleums. *Tetrahedron Lett.* 23:2027–30

Aquino Neto FR, Trendel JM, Restle A, Connan J, Albrecht P. 1983. Occurrence and formation of tricyclic and tetracyclic terpanes in sediments and petroleums. In *Advances in Organic Geochemistry 1981: Proceedings of the 10th International Meeting on Organic Geochemistry*, ed. M Bjoroy, pp. 659–67. Chichester, UK: Wiley

Asara JM, Schweitzer MH, Freimark LM, Phillips M, Cantley LC. 2007. Protein sequences from mastodon and *Tyrannosaurus rex* revealed by mass spectrometry. *Science* 316:280–85

Bacia K, Schwille P, Kurzchalia T. 2005. Sterol structure determines the separation of phases and the curvature of the liquid-ordered phase in model membranes. *PNAS* 102:3272–77

Baker PRS, Armando AM, Campbell JL, Quehenberger O, Dennis EA. 2014. Three-dimensional enhanced lipidomics analysis combining UPLC, differential ion mobility spectrometry, and mass spectrometric separation strategies. *J. Lipid Res.* 55:2432–42

Bednarczyk A, Hernandez TC, Schaeffer P, Adam P, Talbot HM, et al. 2005. 32,35-Anhydrobacteriohopanetetrol: an unusual bacteriohopanepolyol widespread in recent and past environments. *Org. Geochem.* 36:673–77

Bell CD. 2015. Between a rock and a hard place: applications of the "molecular clock" in systematic biology. *Syst. Bot.* 40:6–13

Bertazzo S, Maidment SCR, Kallepitis C, Fearn S, Stevens MM, Xie HN. 2015. Fibres and cellular structures preserved in 75-million-year-old dinosaur specimens. *Nat. Commun.* 6:7352

Blain JC, Szostak JW. 2014. Progress toward synthetic cells. *Annu. Rev. Biochem.* 83:615–40

Blumenberg M, Berndmeyer C, Moros M, Muschalla M, Schmale O, Thiel V. 2013. Bacteriohopanepolyols record stratification, nitrogen fixation and other biogeochemical perturbations in Holocene sediments of the central Baltic Sea. *Biogeosciences* 10:2725–35

Blumenberg M, Kruger M, Nauhaus K, Talbot HM, Oppermann BI, et al. 2006. Biosynthesis of hopanoids by sulfate-reducing bacteria (genus *Desulfovibrio*). *Environ. Microbiol.* 8:1220–27

Bradley AS, Pearson A, Sáenz JP, Marx CJ. 2010. Adenosylhopane: the first intermediate in hopanoid side chain biosynthesis. *Org. Geochem.* 41:1075–81

Brasier MD, Green OR, Jephcoat AP, Kleppe AK, Van Kranendonk MJ, et al. 2002. Questioning the evidence for Earth's oldest fossils. *Nature* 416:76–81

Breslow DK, Weissman JS. 2010. Membranes in balance: mechanisms of sphingolipid homeostasis. *Mol. Cell* 40:267–79

Brocks JJ. 2011. Millimeter-scale concentration gradients of hydrocarbons in Archean shales: live-oil escape or fingerprint of contamination? *Geochim. Cosmochim. Acta* 75:3196–213

Brocks JJ, Banfield J. 2009. Unravelling ancient microbial history with community proteogenomics and lipid geochemistry. *Nat. Rev. Microbiol.* 7:601–9

Brocks JJ, Logan GA, Buick R, Summons RE. 1999. Archean molecular fossils and the early rise of eukaryotes. *Science* 285:1033–36

Brocks JJ, Love GD, Summons RE, Knoll AH, Logan GA, Bowden SA. 2005. Biomarker evidence for green and purple sulphur bacteria in a stratified Palaeoproterozoic sea. *Nature* 437:866–70

Brocks JJ, Summons RE. 2003. Sedimentary hydrocarbons, biomarkers for early life. In *Treatise on Geochemistry*, Vol. 8: *Biogeochemistry*, ed. WH Schlesinger, pp. 63–115. Oxford, UK: Pergamon. 1st ed.

Carroll SB. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134:25–36

Cooke MP, Talbot HM, Wagner T. 2008. Tracking soil organic carbon transport to continental margin sediments using soil-specific hopanoid biomarkers: a case study from the Congo fan (ODP site 1075). *Org. Geochem.* 39:965–71

Crosby CH, Bailey JV, Sharma M. 2014. Fossil evidence of iron-oxidizing chemolithotrophy linked to phosphogenesis in the wake of the Great Oxidation Event. *Geology* 42:1015–18

Cunningham JA, Thomas CW, Bengtson S, Marone F, Stampanoni M, et al. 2012. Experimental taphonomy of giant sulphur bacteria: implications for the interpretation of the embryo-like Ediacaran Doushantuo fossils. *Proc. R. Soc. B* 279:1857–64

Cybulski LE, Ballering J, Moussatova A, Inda ME, Vazquez DB, et al. 2015. Activation of the bacterial thermosensor DesK involves a serine zipper dimerization motif that is modulated by bilayer thickness. *PNAS* 112:6353–58

David LA, Alm EJ. 2011. Rapid evolutionary innovation during an Archaean genetic expansion. *Nature* 469:93–96

Doughty DM, Coleman ML, Hunter RC, Sessions AL, Summons RE, Newman DK. 2011. The RND-family transporter, HpnN, is required for hopanoid localization to the outer membrane of *Rhodopseudomonas palustris* TIE-1. *PNAS* 108:E1045–51

Doughty DM, Dieterle M, Sessions AL, Fischer WW, Newman DK. 2014. Probing the subcellular localization of hopanoid lipids in bacteria using NanoSIMS. *PLOS ONE* 9:e84455

Doughty DM, Hunter RC, Summons RE, Newman DK. 2009. 2-Methylhopanoids are maximally produced in akinetes of *Nostoc punctiforme*: geobiological implications. *Geobiology* 7:524–32

Eickhoff M, Birgel D, Talbot HM, Peckmann J, Kappler A. 2013. Oxidation of Fe(II) leads to increased C-2 methylation of pentacyclic triterpenoids in the anoxygenic phototrophic bacterium *Rhodopseudomonas palustris* strain TIE-1. *Geobiology* 11:268–78

Elvert M, Niemann H. 2008. Occurrence of unusual steroids and hopanoids derived from aerobic methanotrophs at an active marine mud volcano. *Org. Geochem.* 39:167–77

Fahy E, Subramaniam S, Murphy RC, Nishijima M, Raetz CRH, et al. 2009. Update of the LIPID MAPS comprehensive classification system for lipids. *J. Lipid Res.* 50:S9–14

Farrimond P, Talbot HM, Watson DF, Schulz LK, Wilhelms A. 2004. Methylhopanoids: molecular indicators of ancient bacteria and a petroleum correlation tool. *Geochim. Cosmochim. Acta* 68:3873–82

Fischer WW, Fike DA, Johnson JE, Raub TD, Guan YB, et al. 2014. SQUID-SIMS is a useful approach to uncover primary signals in the Archean sulfur cycle. *PNAS* 111:5468–73

Fischer WW, Summons RE, Pearson A. 2005. Targeted genomic detection of biosynthetic pathways: anaerobic production of hopanoid biomarkers by a common sedimentary microbe. *Geobiology* 3:33–40

Förster HJ, Biemann K, Haigh WG, Tattrie NH, Colvin JR. 1973. The structure of novel C_{35} pentacyclic terpenes from *Acetobacter xylinum*. *Biochem. J.* 135:133–43

Freeman KH, Hayes JM, Trendel JM, Albrecht P. 1990. Evidence from carbon isotope measurements for diverse origins of sedimentary hydrocarbons. *Nature* 343:254–56

Articulates criteria for assessing biomarker syngenicity in Archean rocks.

French KL, Hallmann C, Hope JM, Schoon PL, Zumberge JA, et al. 2015. Reappraisal of hydrocarbon biomarkers in Archean rocks. *PNAS* 112:5915–20

Frickey T, Kannenberg E. 2009. Phylogenetic analysis of the triterpene cyclase protein family in prokaryotes and eukaryotes suggests bidirectional lateral gene transfer. *Environ. Microbiol.* 11:1224–41

Frigaard NU, Maresca JA, Yunker CE, Jones AD, Bryant DA. 2004. Genetic manipulation of carotenoid biosynthesis in the green sulfur bacterium *Chlorobium tepidum*. *J. Bacteriol.* 186:5210–20

Hancock JF. 2006. Lipid rafts: contentious only from simplistic standpoints. *Nat. Rev. Mol. Cell Biol.* 7:456–62

Härtner T, Straub KL, Kannenberg E. 2005. Occurrence of hopanoid lipids in anaerobic *Geobacter* species. *FEMS Microbiol. Lett.* 243:59–64

Hays LE, Grice K, Foster CB, Summons RE. 2012. Biomarker and isotopic trends in a Permian-Triassic sedimentary section at Kap Stosch, Greenland. *Org. Geochem.* 43:67–82

Higgins MB, Robinson RS, Husson JM, Carter SJ, Pearson A. 2012. Dominant eukaryotic export production during ocean anoxic events reflects the importance of recycled NH_4^+. *PNAS* 109:2269–74

Jenkyns HC. 2010. Geochemistry of oceanic anoxic events. *Geochem. Geophys. Geosyst.* 11:Q03004

Kalisky T, Blainey P, Quake SR. 2011. Genomic analysis at the single-cell level. *Annu. Rev. Genet.* 45:431–45

Kannenberg E, Blume A, McElhaney RN, Poralla K. 1983. Monolayer and calorimetric studies of phosphatidylcholines containing branched chain fatty acids and of their interactions with cholesterol and with a bacterial hopanoid in model membranes. *Biochim. Biophys. Acta* 733:111–16

Kharbush JJ, Ugalde JA, Hogle SL, Allen EE, Aluwihare LI. 2013. Composite bacterial hopanoids and their microbial producers across oxygen gradients in the water column of the California current. *Appl. Environ. Microbiol.* 79:7491–501

Kiser PD, Golczak M, Maeda A, Palczewski K. 2012. Key enzymes of the retinoid (visual) cycle in vertebrate retina. *Biochim. Biophys. Acta* 1821:137–51

Kleemann G, Alskog G, Berry AM, Hussdanell K. 1994. Lipid composition and nitrogenase activity of symbiotic *Frankia* (*Alnus incana*) in response to different oxygen concentrations. *Protoplasma* 183:107–15

Knani M, Corpe WA, Rohmer M. 1994. Bacterial hopanoids from pink-pigmented facultative methylotrophs (PPFMS) and from green plant surfaces. *Microbiology* 140:2755–59

Knoll AH, Summons RE, Waldbauer JR, Zumberge JE. 2007. The geological succession of primary producers in the oceans. In *Evolution of Primary Producers in the Sea*, ed. PG Falkowski, AH Knoll, pp. 133–63. Amsterdam: Elsevier

Kulkarni G, Busset N, Molinaro A, Gargani D, Chaintreuil C, et al. 2015. Specific hopanoid classes differentially affect free-living and symbiotic states of *Bradyrhizobium diazoefficiens*. *mBio* 6:e01251–15

Kulkarni G, Wu CH, Newman DK. 2013. The general stress response factor EcfG regulates expression of the C-2 hopanoid methylase HpnP in *Rhodopseudomonas palustris* TIE-1. *J. Bacteriol.* 195:2490–98

Kunin V, Goldovsky L, Darzentas N, Ouzounis CA. 2005. The net of life: reconstructing the microbial phylogenetic network. *Genome Res.* 15:954–59

Kuypers MMM, van Breugel Y, Schouten S, Erba E, Sinninghe Damsté JS. 2004. N_2-fixing cyanobacteria supplied nutrient N for Cretaceous oceanic anoxic events. *Geology* 32:853–56

Lingwood D, Simons K. 2010. Lipid rafts as a membrane-organizing principle. *Science* 327:46–50

Lombard J, Lopez-Garcia P, Moreira D. 2012. The early evolution of lipid membranes and the three domains of life. *Nat. Rev. Microbiol.* 10:507–15

Love GD, Grosjean E, Stalvies C, Fike DA, Grotzinger JP, et al. 2009. Fossil steroids record the appearance of Demospongiae during the Cryogenian period. *Nature* 457:718–21

Margheri MC, Allotta G. 1993. Homoacetic fermentation in the cyanobacterium *Nostoc* sp. strain Cc from *Cycas circinalis*. *FEMS Microbiol. Lett.* 111:213–17

Mercier R, Dominguez-Cuevas P, Errington J. 2012. Crucial role for membrane fluidity in proliferation of primitive cells. *Cell Rep.* 1:417–23

Miller WL, Auchus RJ. 2011. The molecular biology, biochemistry, and physiology of human steroidogenesis and its disorders. *Endocr. Rev.* 32:81–151

Moldowan JM, Seifert WK, Gallegos EJ. 1985. Relationship between petroleum composition and depositional environment of petroleum source rocks. *AAPG Bull.* 69:1255–68

Monteiro FM, Pancost RD, Ridgwell A, Donnadieu Y. 2012. Nutrients as the dominant control on the spread of anoxia and euxinia across the Cenomanian-Turonian oceanic anoxic event (OAE2): model-data comparison. *Paleoceanography* 27:PA4209

Nalin R, Putra SR, Domenach AM, Rohmer M, Gourbiere F, Berry AM. 2000. High hopanoid/total lipids ratio in *Frankia* mycelia is not related to the nitrogen status. *Microbiology* 146:3013–19

Neubauer C, Dalleska NF, Cowley ES, Shikuma NJ, Wu CH, et al. 2015. Lipid remodeling in *Rhodopseudomonas palustris* TIE-1 upon loss of hopanoids and hopanoid methylation. *Geobiology* 13:443–53

Noble R, Alexander R, Kagi RI. 1985. The occurrence of bisnorhopane, trisnorhopane and 25-norhopanes as free hydrocarbons in some Australian shales. *Org. Geochem.* 8:171–76

Ottesen EA, Hong JW, Quake SR, Leadbetter JR. 2006. Microfluidic digital PCR enables multigene analysis of individual environmental bacteria. *Science* 314:1464–67

Ourisson G, Albrecht P. 1992. Hopanoids. 1. Geohopanoids: the most abundant natural products on Earth? *Acc. Chem. Res.* 25:398–402

Ourisson G, Rohmer M. 1992. Hopanoids. 2. Biohopanoids: a novel class of bacterial lipids. *Acc. Chem. Res.* 25:403–08

Pace NR. 1997. A molecular view of microbial diversity and the biosphere. *Science* 276:734–40

Pais FSM, Ruy PD, Oliveira G, Coimbra RS. 2014. Assessing the efficiency of multiple sequence alignment programs. *Algorithms Mol. Biol.* 9:4

Pawlowska MM, Butterfield NJ, Brocks JJ. 2013. Lipid taphonomy in the Proterozoic and the effect of microbial mats on biomarker preservation. *Geology* 41:103–6

Pearson A. 2014. Lipidomics for geochemistry. In *Treatise on Geochemistry*, Vol. 12: *Organic Geochemistry*, ed. PG Falkowski, KH Freeman, pp. 291–336. Oxford, UK: Elsevier. 2nd ed.

Pearson A, Leavitt WD, Sáenz JP, Summons RE, Tam MCM, Close HG. 2009. Diversity of hopanoids and squalene-hopene cyclases across a tropical land-sea gradient. *Environ. Microbiol.* 11:1208–23

Pearson A, Page SRF, Jorgenson TL, Fischer WW, Higgins MB. 2007. Novel hopanoid cyclases from the environment. *Environ. Microbiol.* 9:2175–88

Pearson A, Rusch DB. 2009. Distribution of microbial terpenoid lipid cyclases in the global ocean metagenome. *ISME J.* 3:352–63

Pedersen MW, Overballe-Petersen S, Ermini L, Sarkissian CD, Haile J, et al. 2015. Ancient and modern environmental DNA. *Philos. Trans. R. Soc. B* 370:20130383

Peters KE, Moldowan JM. 1991. Effects of source, thermal maturity, and biodegradation on the distribution and isomerization of homohopanes in petroleum. *Org. Geochem.* 17:47–61

Peters KE, Moldowan JM, Walters CC. 2005. *Biomarkers and Isotopes in the Environment and Human History*. Cambridge, UK: Cambridge Univ. Press

Petroff AP, Sim MS, Maslov A, Krupenin M, Rothman DH, Bosak T. 2010. Biophysical basis for the geometry of conical stromatolites. *PNAS* 107:9956–61

Poger D, Mark AE. 2013. The relative effect of sterols and hopanoids on lipid bilayers: when comparable is not identical. *J. Phys. Chem.* 117:16129–40

Poralla K, Kannenberg E, Blume A. 1980. A glycolipid containing hopane isolated from the acidophilic, thermophilic *Bacillus acidocaldarius*, has a cholesterol-like function in membranes. *FEBS Lett.* 113:107–10

Puth K, Hofbauer HF, Sáenz JP, Ernst R. 2015. Homeostatic control of biological membranes by dedicated lipid and membrane packing sensors. *Biol. Chem.* 396:1043–58

Early review of hopanoid discovery and distribution patterns in the rock record.

Rashby SE, Sessions AL, Summons RE, Newman DK. 2007. Biosynthesis of 2-methylbacteriohopanepolyols by an anoxygenic phototroph. *PNAS* 104:15099–104

Discovery that an alphaproteobacterium can produce 2-methylhopanoids in equal abundance as can cyanobacteria.

Rasmussen B, Fletcher IR, Brocks JJ, Kilburn MR. 2008. Reassessing the first appearance of eukaryotes and cyanobacteria. *Nature* 455:1101–4

Ricci JN. 2015. *Constraining the interpretation of 2-methylhopanoids through genetic and phylogenetic methods*. PhD Diss., Div. Biol. Biol. Eng., Calif. Inst. Technol.

Ricci JN, Coleman ML, Welander PV, Sessions AL, Summons RE, et al. 2014. Diverse capacity for 2-methylhopanoid production correlates with a specific ecological niche. *ISME J.* 8:675–84

Identification of a correlation between 2-methylhopanoid production capacity and a particular ecological niche.

Ricci JN, Michel AJ, Newman DK. 2015. Phylogenetic analysis of HpnP reveals the origin of 2-methylhopanoid production in Alphaproteobacteria. *Geobiology* 13:267–77

Rohmer M, Bouvier P, Ourisson G. 1979. Molecular evolution of biomembranes—structural equivalents and phylogenetic precursors of sterols. *PNAS* 76:847–51

A phylogenetic analysis of HpnP indicating that 2-methylhopanoid biosynthesis originated in the Alphaproteobacteria.

Rohmer M, Bouvier-Nave P, Ourisson G. 1984. Distribution of hopanoid triterpenes in prokaryotes. *J. Gen. Microbiol.* 130:1137–50

Rohmer M, Ourisson G. 1976. Structure of bacteriohopanetetrols from *Acetobacter xylinum*. *Tetrahedron Lett.* 40:3633–36

Rush D, Sinninghe Damsté JS, Poulton SW, Thamdrup B, Garside AL, et al. 2014. Anaerobic ammonium-oxidising bacteria: a biological source of the bacteriohopanetetrol stereoisomer in marine sediments. *Geochim. Cosmochim. Acta* 140:50–64

Sackmann E. 1995. Biological membranes—architecture and function. In *Structure and Dynamics of Membranes from Cells to Vesicles*, ed. R Lipowsky, E Sackmann, pp. 1–63. Amsterdam: North-Holland

Sáenz JP, Eglinton TI, Summons RE. 2011a. Abundance and structural diversity of bacteriohopanepolyols in suspended particulate matter along a river to ocean transect. *Org. Geochem.* 42:774–80

Sáenz JP, Grosser D, Bradley AS, Lagny TJ, Lavrynenko O, et al. 2015. Hopanoids as functional analogues of cholesterol in bacterial membranes. *PNAS* 112:11971–76

Sáenz JP, Sezgin E, Schwille P, Simons K. 2012a. Functional convergence of hopanoids and sterols in membrane ordering. *PNAS* 109:14236–40

Sáenz JP, Wakeham SG, Eglinton TI, Summons RE. 2011b. New constraints on the provenance of hopanoids in the marine geologic record: bacteriohopanepolyols in marine suboxic and anoxic environments. *Org. Geochem.* 42:1351–62

Sáenz JP, Waterbury JB, Eglinton TI, Summons RE. 2012b. Hopanoids in marine cyanobacteria: probing their phylogenetic distribution and biological role. *Geobiology* 10:311–19

Schoon PL, Sluijs A, Sinninghe Damsté JS, Schouten S. 2011. Stable carbon isotope patterns of marine biomarker lipids in the Arctic Ocean during Eocene Thermal Maximum 2. *Paleoceanography* 26:PA3215

Schopf JW, Kudryavtsev AB, Walter MR, Van Kranendonk MJ, Williford KH, et al. 2015. Sulfur-cycling fossil bacteria from the 1.8-Ga Duck Creek Formation provide promising evidence of evolution's null hypothesis. *PNAS* 112:2087–92

Schweitzer MH, Zheng WX, Organ CL, Avci R, Suo ZY, et al. 2009. Biomolecular characterization and protein sequences of the Campanian Hadrosaur *B. canadensis*. *Science* 324:626–31

Seckler B, Poralla K. 1986. Characterization and partial purification of squalene-hopene cyclase from *Bacillus acidocaldarius*. *Biochim. Biophys. Acta* 881:356–63

Sharpe HJ, Stevens TJ, Munro S. 2010. A comprehensive comparison of transmembrane domains reveals organelle-specific properties. *Cell* 142:158–69

Silipo A, Vitiello G, Gully D, Sturiale L, Chaintreuil C, et al. 2014. Covalently linked hopanoid-lipid A improves outer-membrane resistance of a *Bradyrhizobium* symbiont of legumes. *Nat. Commun.* 5:5106

Simons K, Ikonen E. 1997. Functional rafts in cell membranes. *Nature* 387:569–72

Sinninghe Damsté JS, Schouten S. 1997. Is there evidence for a substantial contribution of prokaryotic biomass to organic carbon in Phanerozoic carbonaceous sediments? *Org. Geochem.* 26:517–30

Speelman EN, Reichart GJ, de Leeuw JW, Rijpstra WIC, Sinninghe Damsté JS. 2009. Biomarker lipids of the freshwater fern *Azolla* and its fossil counterpart from the Eocene Arctic Ocean. *Org. Geochem.* 40:628–37

Stal LJ. 1991. The metabolic versatility of the mat-building cyanobacteria *Microcoleus chthonoplastes* and *Oscillatoria limosa* and its ecological significance. *Arch. Hydrobiol.* 64:453–67

Summons RE, Jahnke LL, Hope JM, Logan GA. 1999. 2-Methylhopanoids as biomarkers for cyanobacterial oxygenic photosynthesis. *Nature* 400:554–57

Talbot HM, Farrimond P. 2007. Bacterial populations recorded in diverse sedimentary biohopanoid distributions. *Org. Geochem.* 38:1212–25

Talbot HM, Farrimond P, Schaeffer P, Pancost RD. 2005. Bacteriohopanepolyols in hydrothermal vent biogenic silicates. *Org. Geochem.* 36:663–72

Talbot HM, Summons RE, Jahnke LL, Cockell CS, Rohmer M, Farrimond P. 2008. Cyanobacterial bacteriohopanepolyol signatures from cultures and natural environmental settings. *Org. Geochem.* 39:232–63

Talbot HM, Watson DF, Pearson EJ, Farrimond P. 2003. Diverse biohopanoid compositions of non-marine sediments. *Org. Geochem.* 34:1353–71

Tanford C. 1978. Hydrophobic effect and organization of living matter. *Science* 200:1012–18

Thiel V, Blumenberg M, Pape T, Seifert R, Michaelis W. 2003. Unexpected occurrence of hopanoids at gas seeps in the Black Sea. *Org. Geochem.* 34:81–87

van Meer G, Voelker DR, Feigenson GW. 2008. Membrane lipids: where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* 9:112–24

Van Mooy BA, Fredricks HF, Pedler BE, Dyhrman ST, Karl DM, et al. 2009. Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity. *Nature* 458:69–72

van Winden JF, Talbot HM, Kip N, Reichart GJ, Pol A, et al. 2012. Bacteriohopanepolyol signatures as markers for methanotrophic bacteria in peat moss. *Geochim. Cosmochim. Acta* 77:52–61

Vogl K, Bryant DA. 2012. Biosynthesis of the biomarker okenone: χ-ring formation. *Geobiology* 10:205–15

Wacey D, Kilburn MR, Saunders M, Cliff J, Brasier MD. 2011. Microfossils of sulphur-metabolizing cells in 3.4-billion-year-old rocks of Western Australia. *Nat. Geosci.* 4:698–702

Welander PV, Coleman ML, Sessions AL, Summons RE, Newman DK. 2010. Identification of a methylase required for 2-methylhopanoid production and implications for the interpretation of sedimentary hopanes. *PNAS* 107:8537–42

Welander PV, Doughty DM, Wu CH, Mehay S, Summons RE, Newman DK. 2012. Identification and characterization of *Rhodopseudomonas palustris* TIE-1 hopanoid biosynthesis mutants. *Geobiology* 10:163–77

Welander PV, Hunter RC, Zhang LC, Sessions AL, Summons RE, Newman DK. 2009. Hopanoids play a role in membrane integrity and pH homeostasis in *Rhodopseudomonas palustris* TIE-1. *J. Bacteriol.* 191:6145–56

Welander PV, Summons RE. 2012. Discovery, taxonomic distribution, and phenotypic characterization of a gene required for 3-methylhopanoid production. *PNAS* 109:12905–10

Wendt KU, Poralla K, Schulz GE. 1997. Structure and function of a squalene cyclase. *Science* 277:1811–15

Wenk MR. 2010. Lipidomics: new tools and applications. *Cell* 143:888–95 ‘

Whiteside JH, Grice K. 2016. Biomarker records associated with mass extinction events. *Annu. Rev. Earth Planet. Sci.* 44:581–612

Willerslev E, Hansen AJ, Binladen J, Brand TB, Gilbert MTP, et al. 2003. Diverse plant and animal genetic records from Holocene and Pleistocene sediments. *Science* 300:791–95

Williford KH, Ushikubo T, Schopf JW, Lepot K, Kitajima K, Valley JW. 2013. Preservation and detection of microstructural and taxonomic correlations in the carbon isotopic compositions of individual Precambrian microfossils. *Geochim. Cosmochim. Acta* 104:165–82

Woese CR, Fox GE. 1977. Phylogenetic structure of prokaryotic domain—primary kingdoms. *PNAS* 74:5088–90

Wu CH, Kong L, Bialecka-Fornal M, Park S, Thompson AL, et al. 2015. Quantitative hopanoid analysis enables robust pattern detection and comparison between laboratories. *Geobiology* 13:391–407

Xie SC, Pancost RD, Yin HF, Wang HM, Evershed RP. 2005. Two episodes of microbial change coupled with Permo/Triassic faunal mass extinction. *Nature* 434:494–97

Xu YP, Cooke MP, Talbot HM, Simpson MJ. 2009. Bacteriohopanepolyol signatures of bacterial populations in Western Canadian soils. *Org. Geochem.* 40:79–86

Zhang CL, Huang ZY, Li YL, Romanek CS, Mills GL, et al. 2007. Lipid biomarkers, carbon isotopes, and phylogenetic characterization of bacteria in California and Nevada hot springs. *Geomicrobiol. J.* 24:519–34

Zhang YM, Rock CO. 2009. Transcriptional regulation in bacterial membrane lipid synthesis. *J. Lipid Res.* 50:S115–19

Original interpretation of 2-methylhopanoids as biomarkers of oxygenic photosynthesis.

Identification of the C-2 methylase HpnP using genetic analysis of the model organism *Rhodopseudomonas palustris*.

Reports the development of new standards for quantitative hopanoid analysis.

Zhu C, Talbot HM, Wagner T, Pan JM, Pancost RD. 2010. Intense aerobic methane oxidation in the Yangtze Estuary: a record from 35-aminobacteriohopanepolyols in surface sediments. *Org. Geochem.* 41:1056–59

Zhu C, Talbot HM, Wagner T, Pan JM, Pancost RD. 2011. Distribution of hopanoids along a land to sea transect: implications for microbial ecology and the use of hopanoids in environmental studies. *Limnol. Oceanogr.* 56:1850–65

Zundel M, Rohmer M. 1985. Prokaryotic triterpenoids. 3. The biosynthesis of 2β-methylhopanoids and 3β-methylhopanoids of *Methylobacterium organophilum* and *Acetobacter pasteurianus* ssp. *pasteurianus*. *Eur. J. Biochem.* 150:35–39

RELATED RESOURCES

Microbial Diversity Course. Woods Hole, MA: Marine Biological Laboratory. **http://www.mbl.edu/microbialdiversity/**

Newman DK, Gralnick JA. 2005. What genetics offers geobiology. *Rev. Mineral. Geochem.* 59:9–26

Newman DK, Orphan VJ, Reysenbach AL. 2012. Molecular biology's contributions to geobiology. In *Fundamentals of Geobiology*, ed. AH Knoll, DE Canfield, KO Konhauser, pp. 228–49. Chichester, UK: Wiley-Blackwell

Workshop on Molecular Evolution. Woods Hole, MA: Marine Biological Laboratory. **http://www.mbl.edu/education/special-topics-courses/workshop-on-molecular-evolution/**

Annual Review
of Earth and
Planetary Sciences
Volume 44, 2016

Contents

Indexes

Errata

An online log of corrections to *Annual Review of Earth and Planetary Sciences* articles may be found at http://www.annualreviews.org/errata/earth