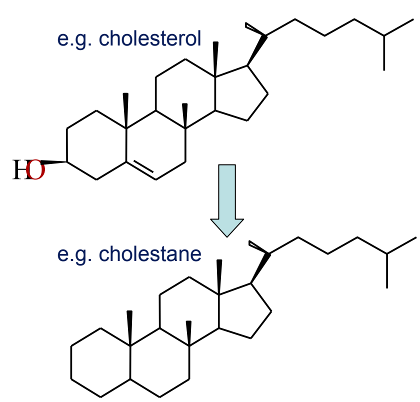
An Introduction to Molecular Markers

- **Key Reading**
- Killops S. and Killops V. (2005) An introduction to Organic Geochemistry, 2nd Edition. Blackwell Scientific. 393 pp.
- **Suggested Reading**
- Volkman J.K., Barrett S.M., Blackburn S.I., Mansour M.P., Sikes E.L. and Gelin F. (1998) Microalgal biomarkers: A review of recent research developments. Org. Geochem. 29, 1163-1179.
- Sinninghe Damste et al., 2004. The Rise of the Rhizosolenid Diatoms. Science. 304, 584-587.
- Coolen et al., 2004. Combined DNA and lipid analyses of sediments reveal changes in Holocene haptophyte and diatom populations in an Antarctic lake. EPSL, 223,225-239.
- Volkman J.K. 2005. Sterols and other triterpenoids: source specificity and evolution of biosynthetic pathways. Org. Geochem. 36, 139-159.

Life, Molecules and the Geological Record

- Life leaves molecular residues (*Chemical Fossils*) as well as visible shapes/objects (*Fossils*) in the sedimentary record.
- These molecular residues, when characterised as specific molecules (*Biomarkers*) by their structures and isotopic content, may give precise indications of their biosynthetic origins in particular organisms, as well as the environmental conditions that the organisms experienced.



Definition of a biomarker (or “molecular marker” or “geochemical fossil”):

“A molecule whose carbon skeleton can unambiguously be linked to that of a known biological precursor compound”

More generally:

“Organic compounds found in sediments which have properties that can be directly related to a known biological precursor”

Biological marker molecules

- Living organisms biosynthesize a very small subset of the billions of molecules that can be assembled in theory from C, H, O, N, S, P etc.
- These molecules can be regarded as biomarkers. Their presence in an environment reflects their synthesis by the parent organisms.
- Some biomolecules are produced only by a certain species or class of organism, and hence indicate the presence or prior existence of those organisms.
- Other biomolecules are produced by many species of organism and are indicative of the general level of biological activity.
- Molecular signatures can comprise the only means to decipher past ecosystems and biological inputs for organisms composed only of soft parts (i.e., leave no morphological imprint).

Molecular Characteristics of biomarkers

- Biomarkers are usually characterized by a high degree of order in their molecular structures, resulting from the specificity of the biosynthetic processes:-
- Small molecule building blocks
- Precise sequence of assembly
- Chirality of carbon centers and stereochemistry of the units
- Distribution of isotopes in the molecule
- Intramolecular characteristics documented by structural identification and molecular isotope measurements.
- Intermolecular variations assessed through compound distributions (e.g. abundance ratios).

Structural uniqueness

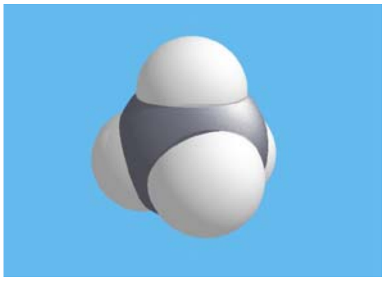
- molecular structure (carbon skeleton)
- stereochemistry

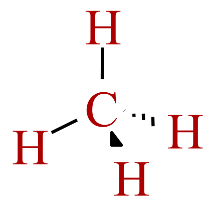
<u>Example</u>: *Only three C_{31} hydrocarbons have been identified in plants (normal- iso- and anteiso-) although there are $>10^9$ possible isomers.*

Distributional uniqueness

- isotopic composition (^{13}C, D/H)
- abundance

Methane, CH_4 - The Smallest Biomarker?





Isotopes:

Carbon: ^{12}C, ^{13}C, ^{14}C

Hydrogen: ^{1}H, ^{2}H, ^{3}H

Universal biomolecular machinery

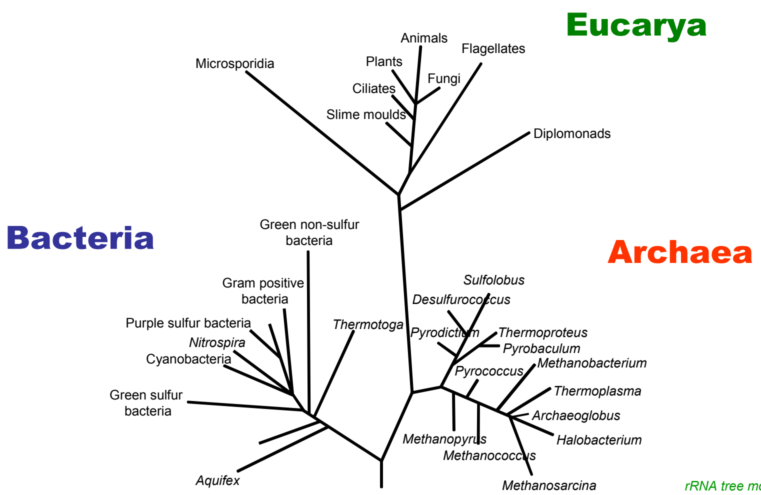
Increasing biomolecular specificity

DNA
↓
RNA
↓
Proteins
↓
Metabolites (sugars, lipids etc.)

Key criteria:

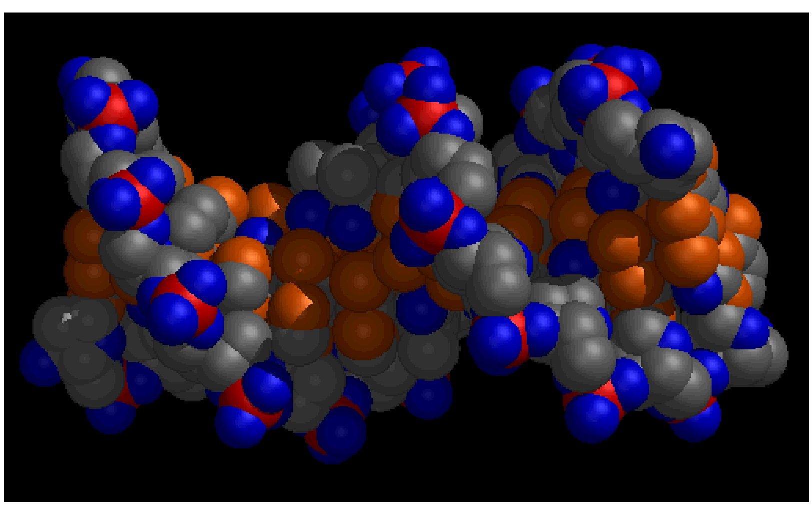
- Information content
- Robustness of molecule
- Ease of detection and analysis (both structural and isotopic)

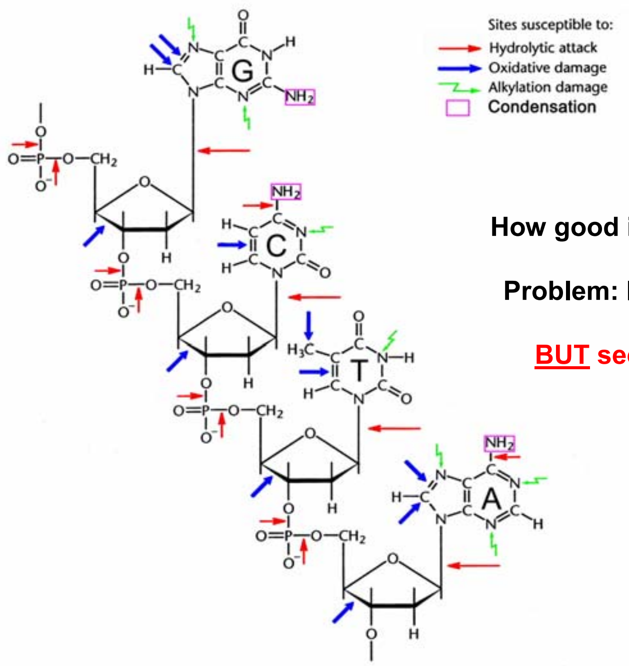
DNA and the phylogenetic tree



rRNA tree modified after Woese

DNA - The Biggest Biomarker!!





How good is DNA as a biomarker?

Problem: It is not preserved well

BUT see Coolen et al., 2004

Lipids

- Lipids present in the water column and in sediments can originate from all three domains of life (i.e., eukaryotes, bacteria, archaea).
- Certain lipids are synthesized by only one domain.
  - Steroids are almost exclusively synthesized by eukaryotes
  - Hopanoids are exclusively synthesized by bacteria
  - Acyclic and cyclic isoprenoid ether lipids are restricted to the archaea.

Occurrence:

- Ubiquitous
- 10-20% of TOC in most organisms
- Extensively studied classes of compounds
- - analytically accessible
- - diagenetically and chemically stable
- - structurally extremely diverse (high potential as "biomarkers")

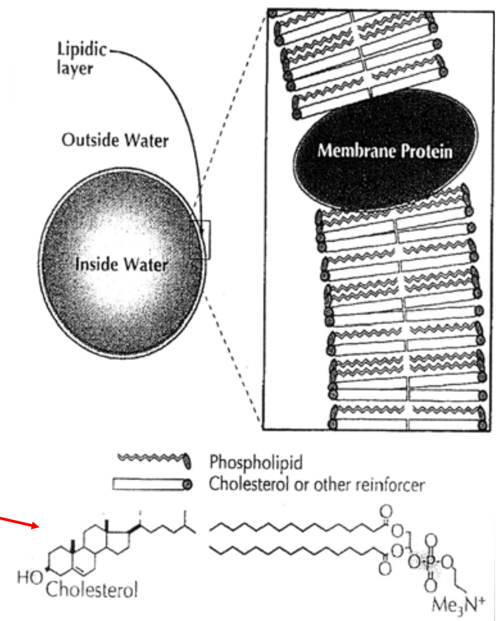
Function:

- - Long-terms energy storage
- - membrane fluidity regulators
- - membrane rigidity/barrier to proton exchange
- - pigments
- - hormones
- - vitamins

All-important cellular membrane: Source of lipids in the environment

Ourisson and Nakatani, 1994

Hopanoid in Bacteria
Acyclic isoprenoid ethers in Archaea



Lipid Structures:

Extremely diverse
Several major compound classes:
- Fatty acids
- Fatty alcohols
- Hydrocarbons
- Terpenoids

Fall into two main groups:
- Polyketide lipids
- Polyisoprene lipids

Occur as "free" compounds or chemically bound (ester or ether linkages) to other biochemical components (e.g., glycerol).

Lipid distributions in plankton

Composition of lipid fraction of diatoms (Clarke and Mazur, 1941)

	%
Uncombined (free) fatty acids	59-82
Combined (bound) fatty acids	1-17
Non-saponifiable (tightly bound) lipids	12-29
Fatty Alcohols	3-7
Hydrocarbons	3-14

Composition of lipid fraction (%) of copepods (Lee et al., 1970)

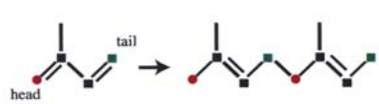
	C. helgolandicus	*G. princeps*
Hydrocarbons	Tr	Tr
Wax esters	37-30	73
Triglycerides	5	9
Polar lipids	14-17	
(free acids, sterols,	50-60	17
phospholipids)		
Total lipid (% dry wt.)	12-15	29

Lipid biosynthesis

Occurs via two main pathways:

1. Polyketide Biosynthesis: The polymerization of acetate; products typically have even carbon numbers.

2. Isoprenoid synthesis: The polymerization of isoprene; products typically have 10, 15, 20 ….. carbon atoms.



Group I. Polyketide lipids

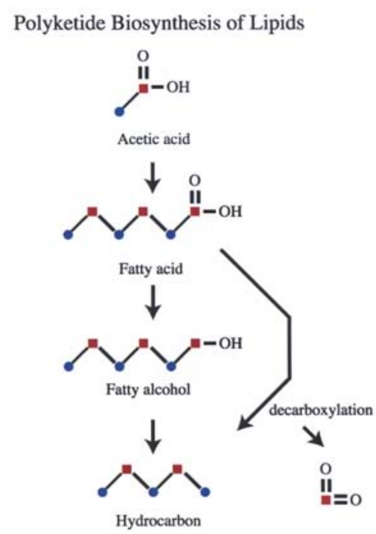
Compounds whose structure is based on repeat units of acetate.

Products usually have an even number of carbons.

Common compound classes:
- Fatty acids
- Fatty alcohols
- Hydrocarbons (*n*-alkanes)

Nomenclature
- *Normal* alkanes and alcohols are denoted: n-C_x where x = number of carbon atoms
- e.g. n-C_{15} = normal pentadecane
- Fatty acids are often unsaturated $C_{x:y}$ where y = number of double bonds
- double bonds are normally *cis*
  Δ denotes the position of unsaturation from the COOH end of the molecule
  - e.g. Oleic acid is $^{9}\Delta\ C_{18:1}$

  ω denotes unsaturation from the methyl end
- Carbon Preference Index (CPI), Odd/Even Predominance (OEP) and Average Chain Length (ACL) used to describe distributions.



Polyketide lipids - Fatty Acids

- Fatty acids are abundant in most organisms (often the most abundant lipid type).
- Sources include bacteria, microalgae, higher plants and marine fauna (e.g., zooplankton).
- Each source has a distinctive profile although some fatty acids are ubiquitous (e.g., C16:0, C18:0).
- Bacteria are a major source of branched fatty acids (iso-, anteiso, mid-chain branched) and can also be a major source of C16:1n-7 and *cis*-vaccenic acid (C18:1n-7).
- Microalgae are a major source of fatty acids in most sedimentary environments.
- Different microalgal inputs can potentially be distinguished based on fatty acid distributions, especially based on # and positions of double bonds.
- Some microalgae contain high concentrations of specific long-chain essential fatty acids (e.g., C20:5n-3; C22:6n-3).

Biosynthesis of Fatty Acids

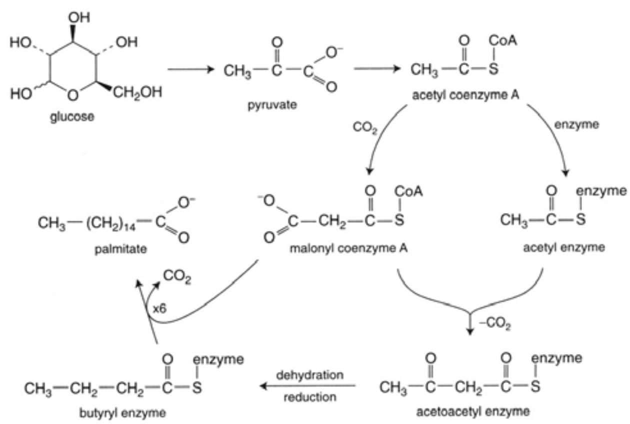


Fig. 2.13 Biosynthesis of saturated fatty acids in plants and animals. Palmitate is formed by successive additions of malonyl coenzyme A to the enzyme-bound chain, with CO_2 being lost at each addition. This results in chain elongation by a $(CH_2)_2$ unit at each step. Details of the formation of butyryl (C_4) from acetyl (C_2) are shown, while the subsequent six further additions, terminating in palmitate, proceed similarly.

Polyketide lipids - Fatty Acids

- In most marine organisms, fatty acids occur predominantly as polar lipids, such as glyco- or phospholipids.
- Free fatty acids are rarely abundant in living organisms, but in sediments they can be the major form due to rapid chemical or enzymatic hydrolysis of polar lipids.
- In the water column and contemporary sediments, intact esterified lipids are usually associated with the indigenous organisms.
- Fatty acid distributions in sediments have been used successfully to characterize bacterial populations.
- A common feature of fatty acid distributions in sediments is the presence of C_{20}-C_{30} saturated straight-chain fatty acids that show a strong predominance of even chain lengths. In many cases these are probably derived from higher plant leaf waxes. However, algae and bacteria can produce these lipids, albeit in small amounts relative to C_{14}-C_{20} acids.

Polyketide lipids - Long-chain alcohols

- Microalgae are not a major source of these lipids in most sediments.
- C_{30-32} alcohols having one or two double bonds are significant constituents of the lipids of marine eustigmatophytes of the genus Nannochloropsis.
- Long-chain diols occur in most marine sediments, and in some cases (e.g., Black Sea Unit I) they can be the major lipids.
- A microalgal source was discovered when C_{30-32} alcohols and diols were identified in marine eustigmatophytes from the genus nannochloropsis (although the distn differed significantly from that in sediments).
- It is suggested that these diols are the building blocks of novel aliphatic biopolymers produced by these microalgae (see below).
- Most studies of mid-chain diols report presence of C_{30} and C_{32} saturated constituents having a predominance of 1,15 isomers.

C30-C32 alkyl diols and unsaturated alcohols in microalgae of the class Eustigmatophyceae

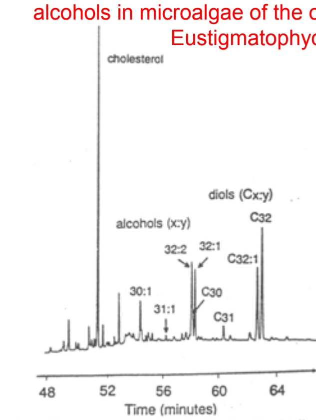


Fig. 1. Partial gas chromatogram showing the distribution of higher molecular weight components in *N. oculata* (strain CS-179) after acid hydrolysis of the extract and derivatization with BSTFA. Alcohols are designated *x*:*y* and diols $C_{x:y}$ where *x* is the number of carbon atoms and *y* is the number of double bonds. Note that the C_{30} diol elutes on the downslope of the peak due to the 32:2 alcohol. Quantitative data are shown in Table 2.

Polyketide lipids - Hydrocarbons

- Biogenic alkanes and alkenes are a common feature of the hydrocarbon distributions in sediments.
- Many microalgae contain the highly unsaturated alkenene n-$C_{21:6}$ (n-heneicosa-3,6,9,12,15,18-hexaene) formed by decarboxylation of the $C_{22:6n-3}$ fatty acid. However, this compound is rarely found in sediments, probably because it is rapidly degraded.
- There are several reports of shorter-chain n-C_{15}, n-C_{17} and n-C_{19} alkanes (and monounsaturated alkenes) in algae. The n-C_{17} alkane is common in contemporary sediments.
- Several microalgae contain very long chain alkenes, including $C_{31:2}$, $C_{33:3}$ and $C_{33:4}$, and $C_{37:2}$ and $C_{37:3}$ (also C_{38} counterparts).

Generalized Lipid Distributions

	Acids	Alcohols	Hydrocarbons
Phytoplankton	even/odd CPI 16:0, 16:1 18:0, 18:1	even/odd CPI	odd/even nC17, nC18
Bacteria	iso + anteiso ?		CPI = 1 nC13-nC30 nC17-nC20
Zooplankton	same as phyto		same as phyto
Higher plants	even/odd CPI max C28-C30	C28, C30, C32	odd/even max C29-C31

- Lipids in higher plants mainly associated with leaf cuticles ("waxes")
- Serve as physical protection

Esterified lipids

Diglyerides and Triglycerides ("fats")

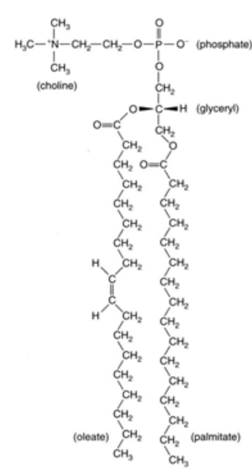
- Esters of glycerol + fatty acids
- formed by condensation (-H_2O) reactions
- fatty acids usually straight-chain with various levels of unsaturation

Wax esters

- Esters comprising of a fatty acid + fatty alcohol

Phospholipids

- Fatty acid + phosphoric acid + glycerol (+ basic nitrogen)



(a) lecithin

Phospholipid

Polyketide lipids - Hydroxy fatty acids

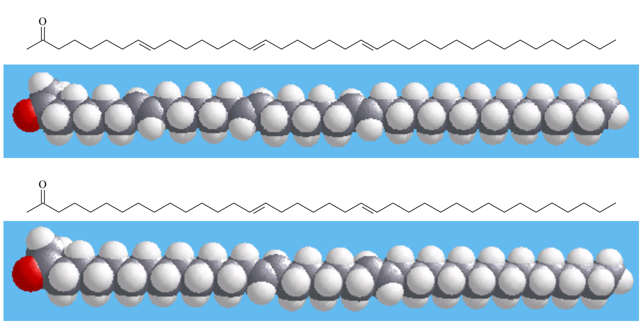
- A wide range of hydroxyl fatty acids has been found in sediments, these compounds have received relatively little attention from organic geochemists.
- These compounds can be separated into different categories based on the # and position of the hydroxyl groups.
- Aliphatic α- and β-monohydroxyfatty acids occur in a wide range of organisms.
- α-hydroxy fatty acids are intermediates in fatty acid biosynthesis in yeasts.
- Bacterially-derived β-hydroxy fatty acids are found in many recent sediments. The carbon # range is typically from C_{10}-C_{20}, typical of carbon # distribution for lipopolysaccharide cell walls of gram negative bacteria. Bacteria also contribute significant amounts of *iso* and *anteiso* branched C_{12}-C_{18} β-hydroxy fatty acids.
- Higher plant cutin and suberin can also be a significant source of esterified C_{16}-C_{22} α-, β-, and ω-monohydroxy fatty acids.
- Recent work suggests microalgae are also a potential source of monohydroxy fatty acids.
- C_{30}-C_{34} mid-chain hydroxyl fatty acids were identified in hydrolyzed extracts of marine eustigmatophytes of the genus Nannochloropsis.
- C_{22}-C_{26} saturated and monounsaturated a-hydroxyfatty acids have also been found as major lipid components of the cell wall of several marine chlorophytes.

Polyketide lipids - Long-chain ketones, esters

- Long-chain unsaturated ketones (alkenones) have been identified in several species of haptophytes, esp. the widely distributed coccolithophorids *Emiliania huxleyi* and *Gephyrocapsa oceanica*.
- These compounds will be the focus of a separate lecture.

37:3

37:2



E. huxleyi

Group II. Polyisoprenoid lipids

Biosynthetically related to the polymerization of isoprene (C5)

- dimer - monoterpene (C_{10})
- e.g. essential oils

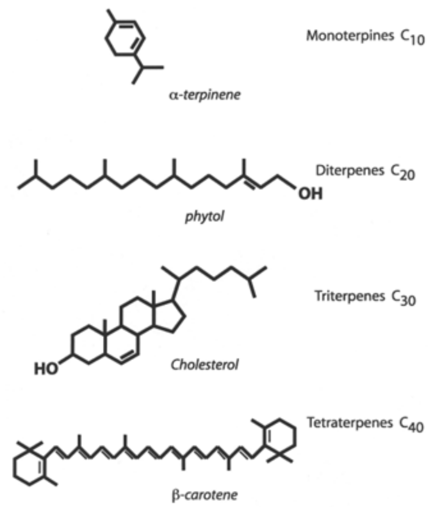
- trimer - sesquiterpenes (C_{15})
- e.g. farnesol, abeitic acid

- tetramer - diterpenes (C_{20})
- phytol

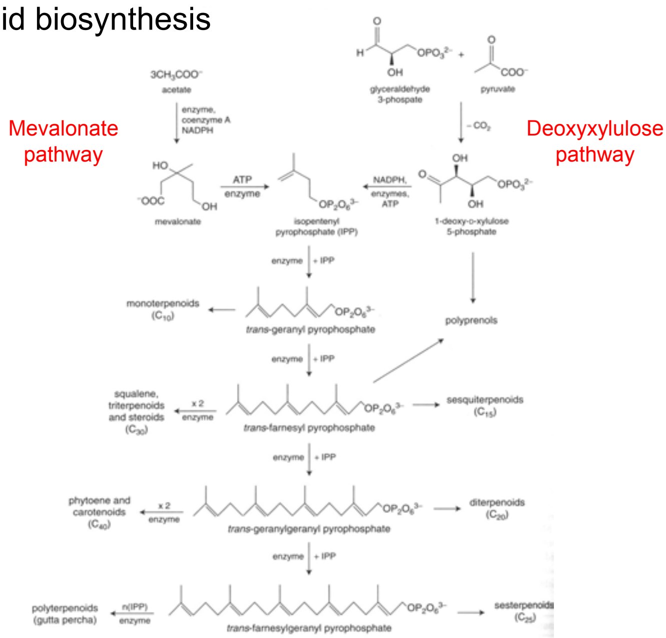
- hexamer - triterpenes (C_{30})
- steroids, hopanoids

- octamer - tetraterpenes (C_{40})
- carotenoids, ether lipids

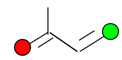
- polymer - polyterpenes
- e.g. natural resins (polycadinene)



Isoprenoid lipid biosynthesis



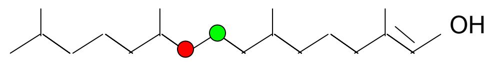
Isoprenoid lipids - Configurations



isoprene

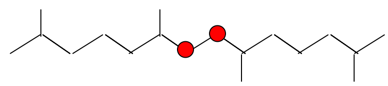
Regular isoprenoids

- head-to-tail
  - e.g. phytol

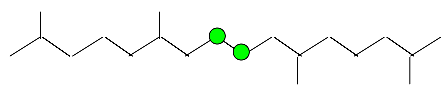


Irregular isoprenoids

- head-to-head



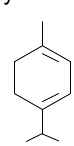
- tail-to-tail



Isoprenoid lipids

Monoterpenes (C10)

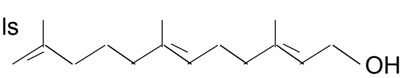
- Abundant in both higher plants and algae
- not extensively used as biomarkers (very volatile, so not well preserved in sediments)



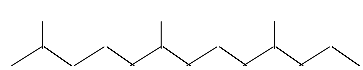
α-terpinene

Sesquiterpenes (C15)

- farnesol
- esterified to bacterial chlorophylls

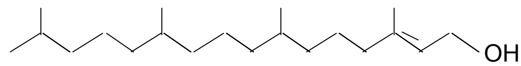


- ancient analogue - "farnesane"



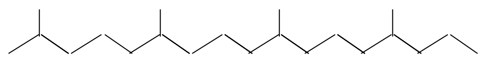
Isoprenoid lipids

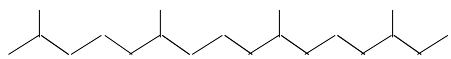
Diterpenes (C20)
Acyclic diterpenoids
- phytol: trans 3,7(R),11(R),15-tetramethylhexadec-2-ene-1-ol



sources:
- esterified to chlorophyll-a,b
- phytanylethers (archaebacteria)

ancient analogues
- phytane
- pristane

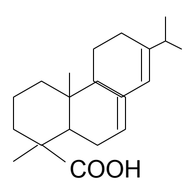




origin of pristane
- chlorophyll ?
- tocopherol (vitamin e)?

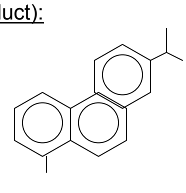
Isoprenoid lipids

Cyclic diterpenoids
- abietic acid



- from gymnosperm (conifer) resin/gum
- not widely found in marine algae
- therefore, excellent "biomarker" for higher plant input

<u>ancient analogue (also combustion product):</u>
- retene



Highly branched isoprenoid alkenes

- Highly branched unsaturated C20, C25, C30 alkenes (containing between 3 and 6 double bonds) are observed in most marine sediments. These compounds appear to derive exclusively from diatoms, although the function and bioactivity of these compounds is yet to be established.

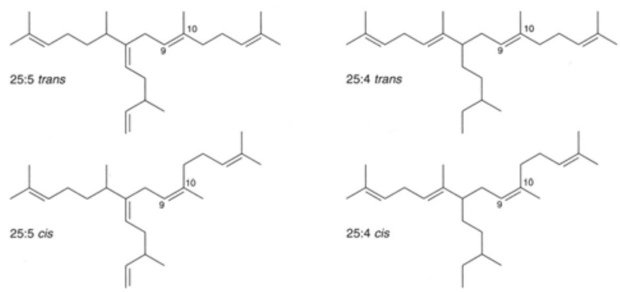


Fig. 2.18 Examples of highly branched sesterterpenes isolated from planktonic diatom cultures (*trans* and *cis* refer to configuration at double bond between C-9 and C-10; after Belt et al. 2001).

Isoprenoid lipids

Cyclic Triterpenoids (C30)

- Squalene - main biosynthetic precursor to cyclic triterpenes.
- An irregular isoprenoid (tail-to-tail)

- *Pentacyclic triterpenoids*
- oleanane type
- ursane type
- lupane type
- hopane type
- gammacerane type
- arborane type
- primary sources: bacteria, higher plants

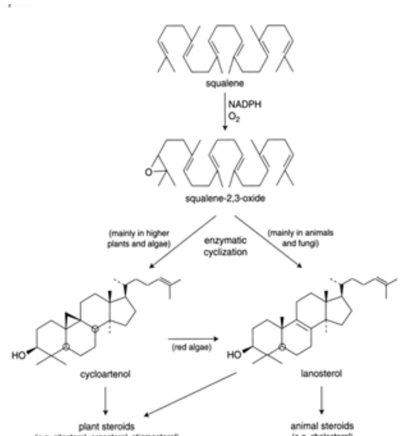


Fig. 2.21 Steroid biosynthesis from squalene (after Harwood & Russell 1984).

Pentacyclic triterpenoids

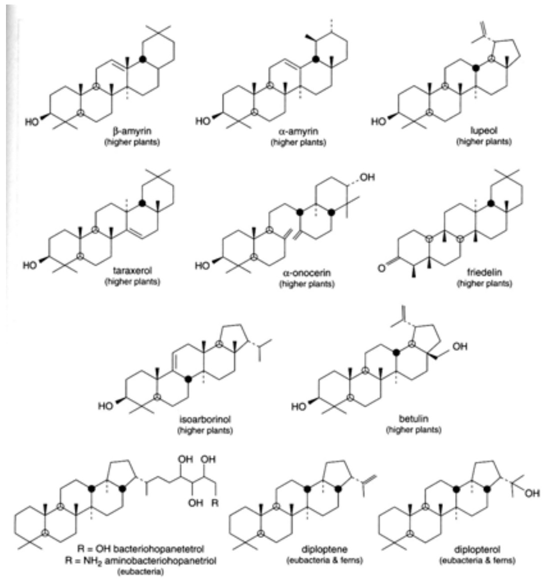


Fig. 2.19 Some geochemically important polycyclic triterpenoids and their major sources.

Carbon numbering of steroids and hopanoids

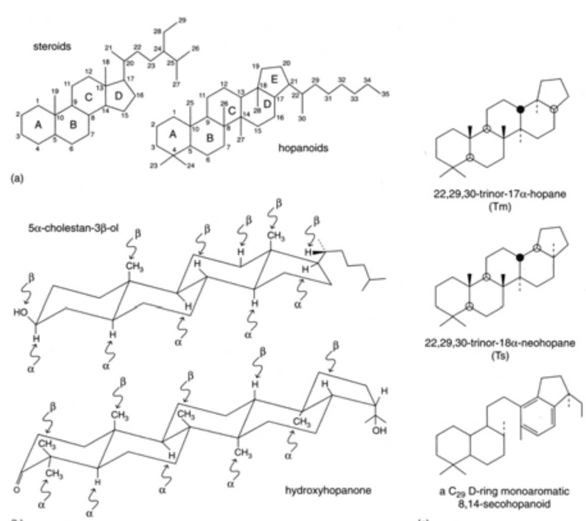


Fig. 2.20 (a) Ring numbering conventions for steroids and hopanoids; (b) examples of 'all-chair' conformations (with *trans* ring junctions) for steroids and hopanoids; (c) examples of application of hopanoidal nomenclature system. (Ts may also be called 17-methyl-18,22,29,30-tetranor-17α,18α-hopane.)

Isoprenoid lipids - Tetracyclic triterpenoids (Sterols)

Numbering and nomenclature:

- -, - below (dashed arrow) and above (bold arrow) the ring
- *R/S* stereochemistry at a ring juncture and in side chain
- sterol/stenol - unsat'd alcohol
- stanol - sat'd alcohol
- sterene - unsat'd alkene
- sterane - sat'd alkane

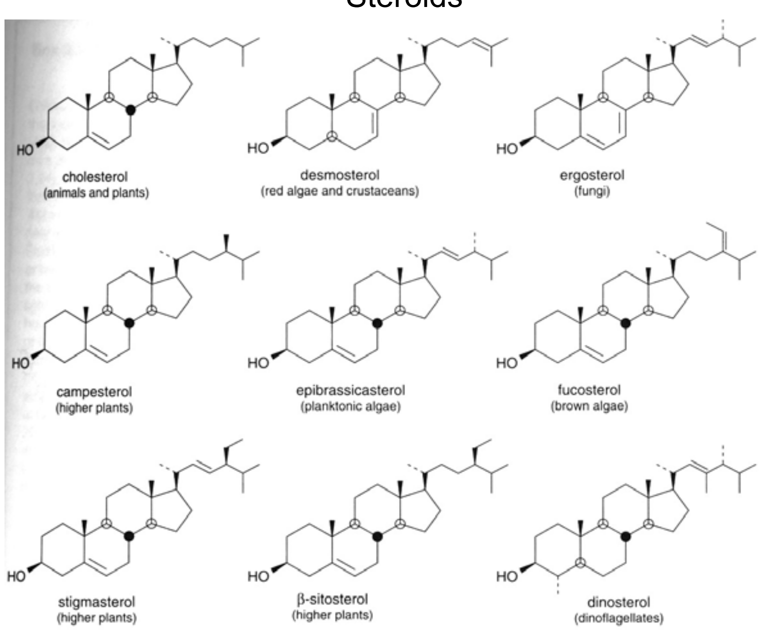
Occurrence:

- very widely distributed in plants and animals
- As a rule, bacteria do not make sterols

Structure:

- Mainly C27, C28, C29 (also C26,C30)
- Basic steroid skeleton is modified through oxidation and alkylation
- Hundreds of natural products based on this skeleton have been identified.
- cholesterol C27 (universally distributed)
- β-sitosterol C29 (higher plants)
- brassicasterol (diatoms)
- dinosterol C30 (dinoflagellates)
- fucosterol (brown algae)

Steroids



Sterols (continued)

- A great diversity of sterols are found in microalgae. Distributions range from the predominance of a single sterol, such as cholesterol in marine eustigmatophytes and 24-methylcholesta-5,22-dien-3b-ol in some diatoms and haptophytes to mixtures of 10 or more 4-methyl and desmethylsterols.
- Some sterols are widely distributed but others are chemotaxanomic markers.
- The diatoms display considerable diversity in sterol composition, and given the importance of diatoms as a source of organic matter in marine systems it is not surprising that sediments display complex and varying sterol distributions.
- Sterols derived from dinoflagellates are often major constituents of the sterol distributions.
- The sterol composition of dinoflagellates is dominated by 4a-methyl sterols, including dinosterol (4a,23,24-trimethyl-5a-cholest-22E-en-3b-ol) – often used as an indicator of dinoflagellate inputs to sediments.
- Sterols with a fully saturated ring system (5a(H)-stanols) often occur in dinoflagellates but are not common in other marine microalgae. Hence dinos are the major direct source of stanols to marine sediments, supplementing those formed by bacterial reduction of stenols.

Other Steroids

- *Related compounds:*
- Steroidal ketones – primarily intermediates in the microbially or chemically mediated degradation in sediments of stenols to sterenes. A direct biological source is also possible.
- Steroidal diols – one species of the genus Pavlova (Haptophyta) contains novel 3,4-dihydroxy-4a-methyl sterols. Source specificity not yet known.

Tetraterpenoids (C40)

- **A. Ether lipids**
- <u>Occurrence:</u>
- Archaea
    - methanogens
    - thermoacidophiles
    - extreme halophiles
    - eurythermal archaeota
- <u>Function:</u>
    - Membrane rigidifiers
- <u>Structure:</u>
- unusual linkage type (mainly head-to-head)
- Ancient analogues:
- head-to-head acyclic isoprenoid alkanes.

- Focus of a separate lecture

Tetraterpenoids - B. Carotenoid pigments

- <u>Occurrence:</u>
- Universally distributed in all photosynthetic organisms
- <u>Function:</u>
- Accessory pigments, antioxidants
- <u>Structure:</u>
- Many different structures (>100 identified to date)
- Bacillophyceae
    - fucoxanthin
    - diadinoxanthin
    - diatoxanthin
    - β-carotene
- Dinophyceae
    - peridinin
- <u>Ancient analogue:</u>
    - β-carotane

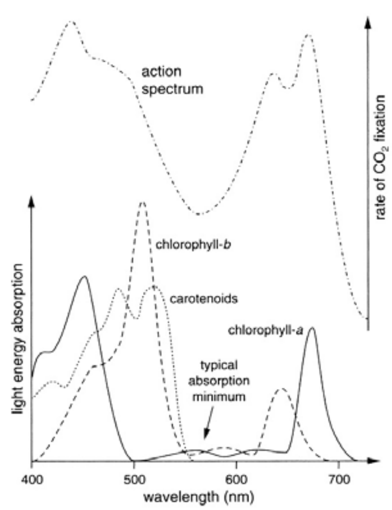
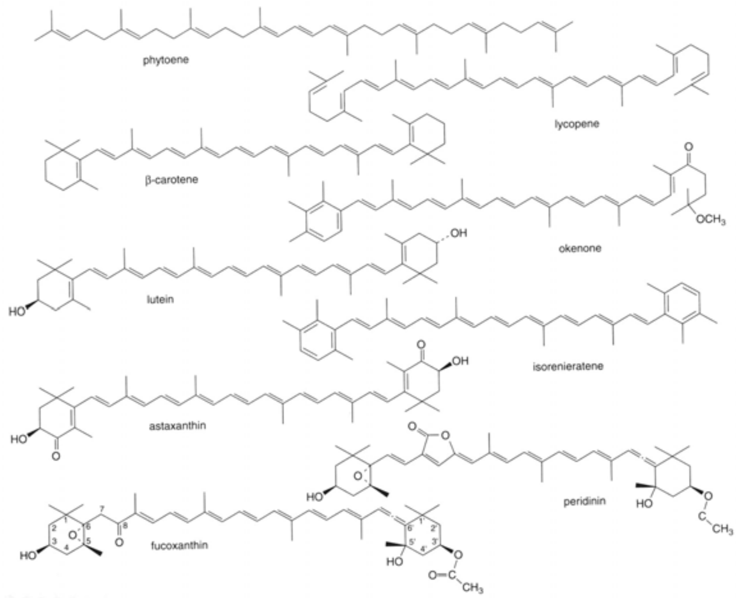


Fig. 2.24 Light absorption characteristics of some photosynthetic pigments and their relationship with utilization of light energy during photosynthesis (action spectrum).

Carotenoids



Chlorophylls

Occurrence:
- Universally distributed in all photosynthetic organisms

Function:
- Used for photosynthesis
- h + chl+ CO_2 + H_2O chl* + O_2 chl + energy + $(CH_2O)n$

Structure:
- All are tetrapyrroles
- Chl-a,b,c1,2,3; a1+a2; b1+b2, d,e, - oxygenic photosynthetic organisms
- bchl-a,b,c,d,e - bacteriochlorophylls

Abundance:
- Ratio of carbon/chl = 60 for phytoplankton

Ancient analogue:
- Porphyrins were the first molecules to be recognized in ancient sediments and petroleum as of biological origin - structurally related to chlorophylls (Treibs, circa1934)

chlorophyll	structure	R^1	R^2	R^3	R^4
chlorophyll-*a*	I	$-CH_2{=}CH_2$	$-CH_3$	$-CH_2-CH_3$	phytyl
chlorophyll-*b*	I	$-CH_2{=}CH_2$	$-C(=O)H$	$-CH_2-CH_3$	phytyl
chlorophyll-a_2	I	$-CH_2{=}CH_2$	$-CH_3$	$-CH_2{=}CH_2$	phytyl
chlorophyll-b_2	I	$-CH_2{=}CH_2$	$-C(=O)H$	$-CH_2{=}CH_2$	phytyl
chlorophyll-c_1	II	$-CH_2{=}CH_2$	$-CH_3$	$-CH_2-CH_3$	H
chlorophyll-c_2	II	$-CH_2{=}CH_2$	$-CH_3$	$-CH{=}CH_2$	H
chlorophyll-*d*	I	$-C(=O)H$	$-CH_3$	$-CH_2-CH_3$	phytyl
bacteriochlorophyll-*a*	III	$-C(=O)CH_3$	$-CH_3$	$-CH_2-CH_3$	phytyl, farnesyl or geranylgeranyl
bacteriochlorophyll-*b*	III	$-C(=O)CH_3$	$-CH_3$	$=CH-CH_3$	phytyl, farnesyl or geranylgeranyl
bacteriochlorophyll-*e*	IV	$-C(H)(OH)CH_3$	$-C(=O)H$	$-CH_2-CH_3$, $-CH_2-CH_2-CH_3$ or $-CH_2-CH(CH_3)-CH_3$	farnesyl

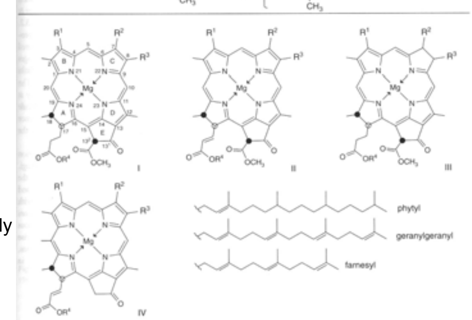
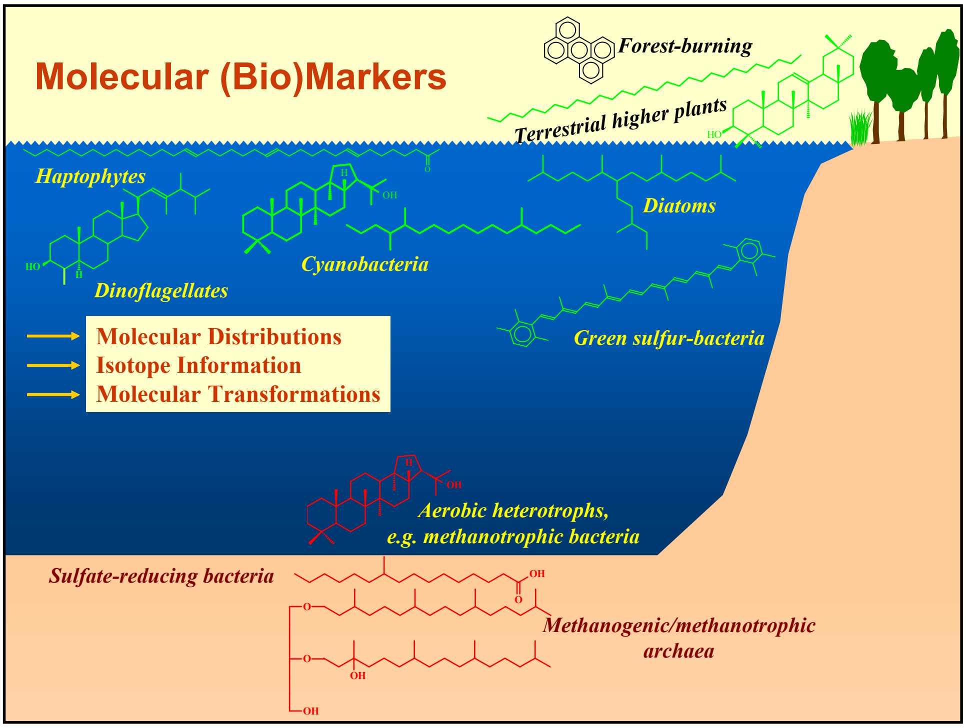


Fig. 2.27 Some geochemically important chlorophylls (ring numbering scheme shown). The arrows from N atoms to Mg^{2+} ions denote dative bonds (see Section 2.1.2).



Biomarker Properties

We can measure and utilise their:

- Precise molecular structures, including stereochemistry
- Relative and absolute amounts
- Isotopic composition – C, H, N etc

Biomarker Research Areas

- Petroleum Exploration
- Environments & Ecology, including anthropogenic effects
- Palaeoenvironment Reconstruction & Climate Change,
  - e.g. Palaeooceanography - Catastrophic events
- Exploration of Biosphere – especially the Microbial World
  - e.g. Extreme environments, deep biosphere etc.
- Evolution of the Biosphere, Origin of Life & Archaean Studies
- Meteoritics, Exobiology & Planetary Studies
  - e.g. Moon, Mars
- Archaeology

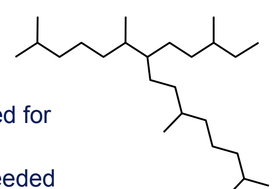
State of the Art: Integration of Lipid Biomarkers & Molecular Biology

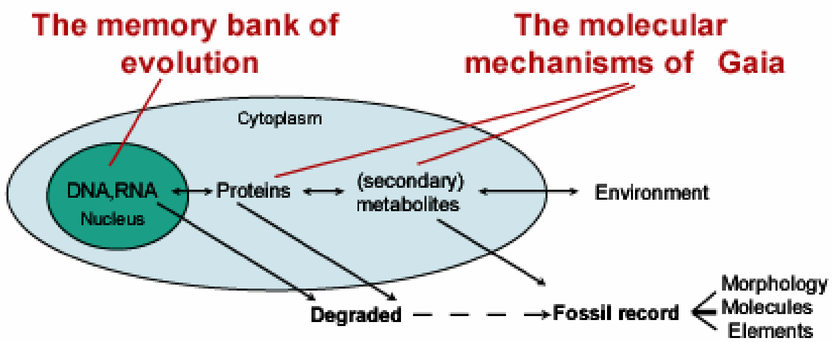
DNA and biomarker chemotaxonomy

RNA and DNA sequence analysis of Groups of living organisms enables :

- Evolutionary Trees to be constructed.
- Specific Biochemical Pathways to be investigated for their distributions in the Phylogenetic Group.
  - E.g. Identifying the gene for the enzymes needed to synthesise particular biomarkers, such as the highly branched isoprenoid (HBI) compounds.

C_{25} HBI





DNA (Genotype) does not interact with environment directly

- Preservation potential of DNA/RNA is very low
- Instead: study living and recent organisms and make phylogenetic tree.

Secondary metabolites (phenotype) interact with environment

- Check contemporary DNA tree for key genes controlling biosynthetic pathways
- Links of secondary metabolites to the environment inform us about their function and can lead to selection of new biomarker proxies.
- Search fossil record for times of first appearance of the biomarkers

DNA and biomarker chemotaxonomy:

When did the Rhizosolenid Diatoms Evolve?

(Sinninghe Damste et al. 2004 Science,304, 584)

- Rhizosolenid diatoms are a very successful Group of marine diatoms. They currently fix around half of the CO_2 flux in the oceans.
- They are the only Group of diatoms to make the HBI Biomarkers
  - 150 Diatom Species analysed for Molecular Phylogeny and HBI Production

DNA and biomarker chemotaxonomy:

When did the Rhizosolenid Diatoms Evolve?

- Genomic analyses of living diatoms gives the Phylogenetic Tree in which the deepest branching points for the Rhizosolenids can be seen.
- Analysis for HBI biomarkers of 81 well-dated petroleums and 700 ancient sediments, going back 0.7 Ma., reveals the first appearance of the HBI at 91.5 Ma in the Upper Turonian.
- So we can date the first appearance of the Rhizosolenids to 91.5Ma, based on the first occurrence of the HBI biomarkers.

Recovery of Fossil DNA from aquatic sediments

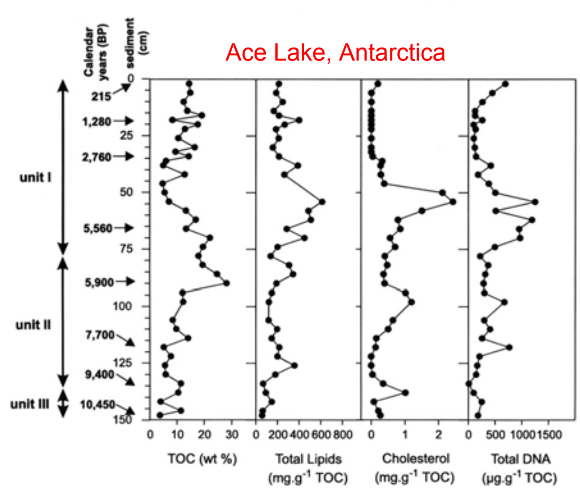


Fig. 3. Depth profiles of total organic carbon (TOC) contents (weight percent), total lipid contents (milligrams per gram TOC), cholesterol as a generic biomarker for overall algal productivity (milligrams per gram of TOC), and extractable DNA (micrograms per gram of TOC). Note that DNA was extractable from even the deepest sediment.

Coolen et al. 2004 EPSL 223, 225-239

Recovery of Fossil DNA from aquatic sediments

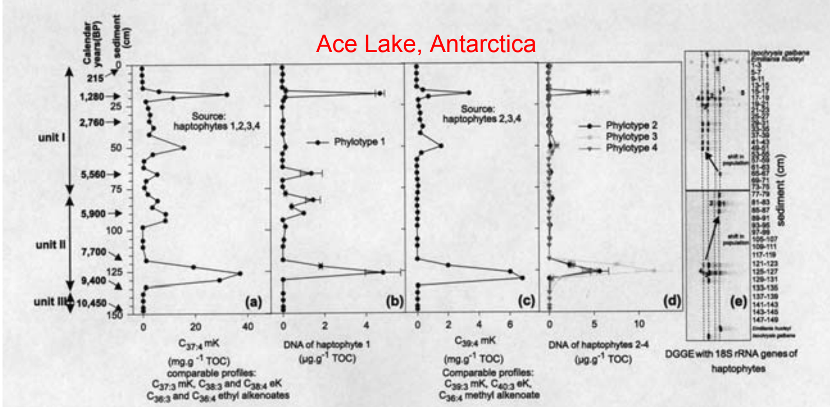


Fig. 4. Stratigraphy of biomarkers of haptophytes (18S rDNA, alkenones [mK, methyl ketone; eK, ethyl ketone], and alkenoates) recovered from the Holocene Ace Lake sediments. (a) Quantity of C_{37}:mK (mg g^{-1} TOC). Comparable profiles were found for $C_{37:3}$ mK, $C_{38:3}$ and $C_{38:4}$ eK as well as C_{36} ethyl alkanoates; (b) Quantity of 18S rDNA of haptophyte 1 (phylotype 1) (µg g^{-1} TOC); (c) quantity of $C_{39:4}$ mK (mg g^{-1} TOC). Comparable profiles were found for $C_{39:3}$ mK, $C_{40:3}$ eK, and a C_{36} methyl alkanoate; (d) quantity of 18S rDNA of haptophytes phylotypes 2–4 (µg g^{-1} sediment); (e) separation of PCR-amplified 18S rDNA of haptophytes from the Holocene sediment layers by DGGE resulted in six fragments with unique melting positions (haptophytes 1–6 numbered in the gel). Phylotypes 5 and 6 (e) were rare and therefore not used for quantitative stratigraphic analysis. The dashed vertical lines indicate identical melting behaviour of phylotypes 1–4 throughout the gel. For comparative analysis, PCR products of reference strains of the haptophytes *I. galbana* CCMP1323 and *E. huxleyi* str. L were separated by DGGE along with the sediment samples. The arrows in the DGGE indicate shifts in the haptophyte populations. Unit III: freshwater lacustrine period. Unit II: due to the rising sea level caused by Holocene deglaciation, marine waters with haptophytes and diatoms entered Ace Lake. Unit I: present-day meromictic saline lacustrine system with euxinic conditions as indicated by molecular remains of green sulfur bacteria (Coolen et al., unpublished results).

Coolen et al. 2004 EPSL 223, 225-239

Recovery of Fossil DNA from aquatic sediments

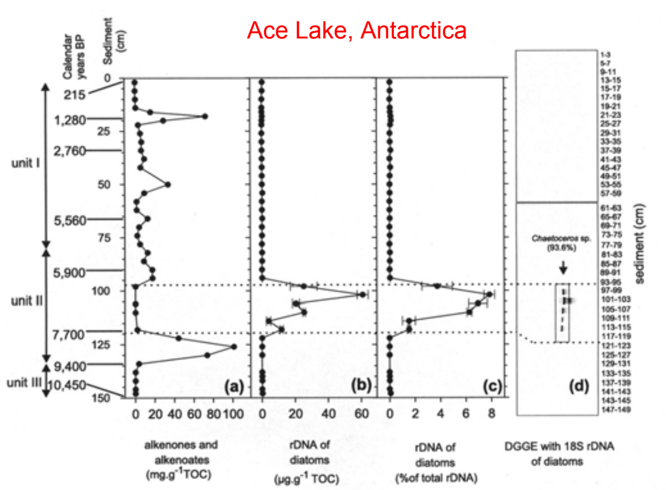
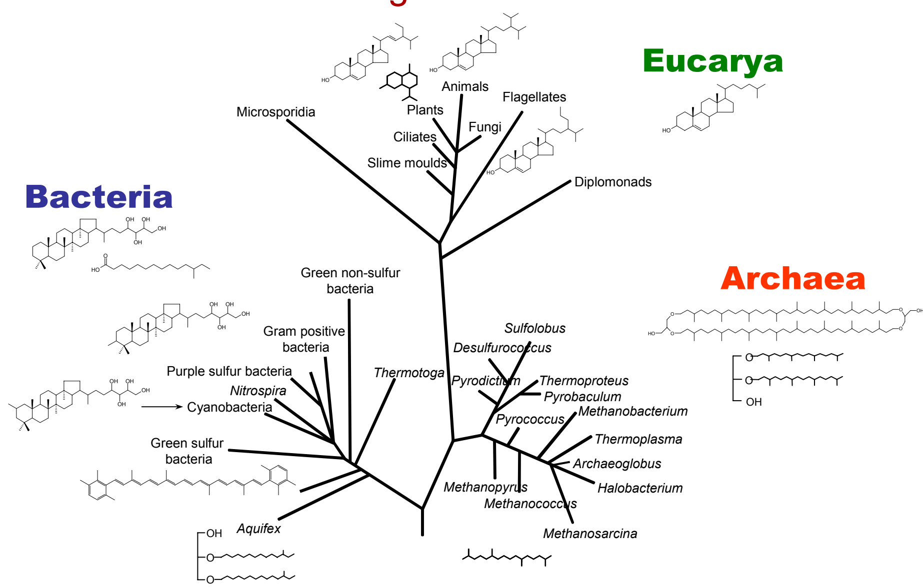


Fig. 6. Abundances of total alkenones and alkenoates (milligrams per gram of TOC) and DNA of diatoms expressed as micrograms per gram of TOC and of total community DNA. DGGE results are shown illustrating the occurrence of high contents of DNA from a diatom related to *Chaetoceros* sp. in the core section deposited after 7700 years BP. Note that this diatom did not become abundant until after the haptophyte populations waned.

Coolen et al. 2004 EPSL 223, 225-239

Summary:
The Biological Precursors…

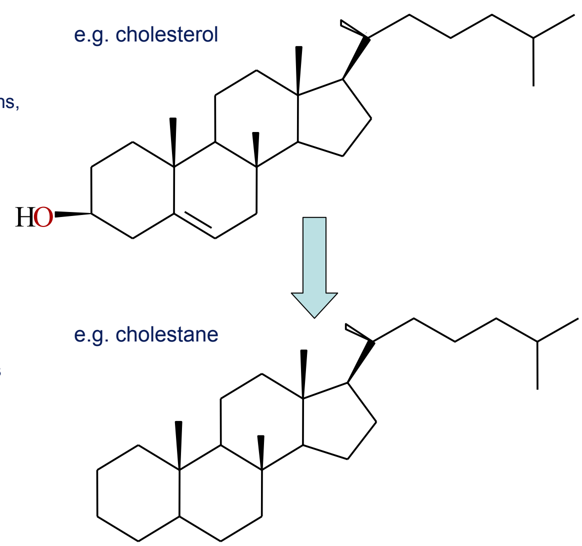


rRNA tree modified after Woese

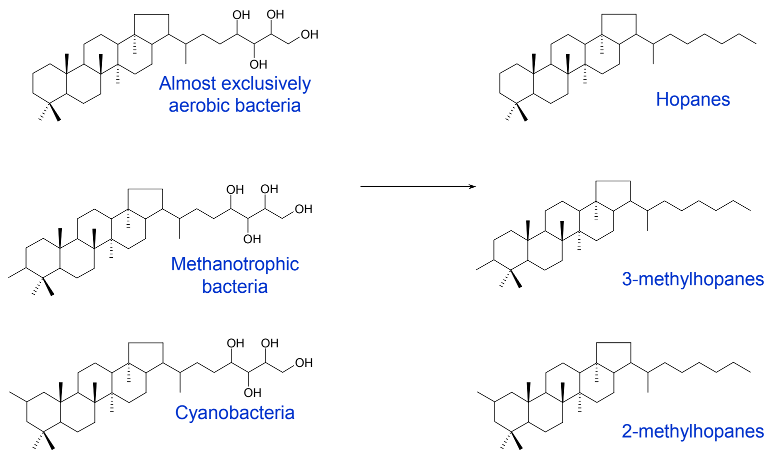
Biomarker Molecules in the Biosphere & Geosphere

Biosphere
Organisms synthesise molecules with highly ordered hydrocarbon skeletons, often carrying several functional groups containing O,N atoms etc.

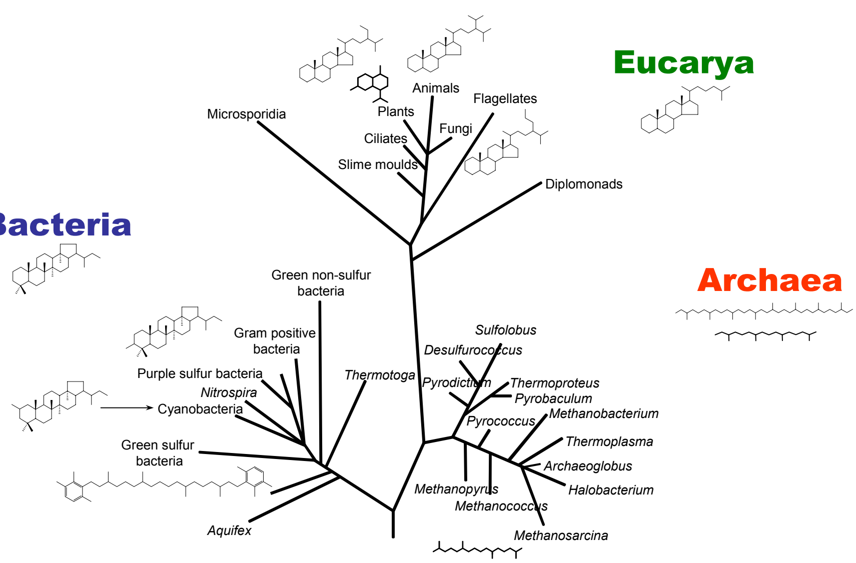
Geosphere
Sedimentary compounds often defunctionalised as parent hydrocarbons with same carbon skeleton.



Hopanoids - The 'Most abundant natural products on Earth?'



…and their Molecular Fossils



rRNA tree modified after Woese

The Microbial Record In The Geosphere