6. Terpenes: C5 to C20

RA Macahig
FM Dayrit

3-(R)-MVA
Introduction

• Terpenes make up very prominent and characteristic group of plant secondary metabolites. Terpene metabolites range from volatile compounds with 10 carbons to colored polyenes with 40 carbons.

• The word “terpene” comes from turpentine, the yellow to brown thick oleoresin which is obtained as an exudate from the terebinth tree (Pistacia terebinthus).

• Terpenes are historically, culturally and economically important:
  • oleoresins, such as pine and eucalyptus oils; rubber (gutta percha)
  • distillates of the resin yield solvents and thinners
  • “essential oils” and perfumes, which are extracted from flowers and leaves by pressing, alcohol extraction or steam distillation
  • drugs and steroids
Introduction

Some characteristic terpenes:

**Camphor**: monoterpane from *Cinamomum camphora*.

**Taxol**: antitumor diterpene from Pacific yew, *Taxus* species

**Cholesterol**: steroid originally isolated from gallstones; component of all cell membranes

**β-Carotene**: C40 terpene which is synthesized in the chloroplast; important plant pigment; believed to be one of the important natural anti-oxidants.
**Introduction**

- In the late 19th century, Otto Wallach noted that upon chemical degradation, many of the products obtained had chemical formulas which were in multiples of 5 carbons. In the 1860s, these C5 units were called “isoprene” units. This is the basis of the “**isoprene rule**” which was formulated by Leopold Ruzicka. Isoprene represents the basic skeletal structure of the C5 unit.

The most prolific producer of isoprene-type polymers is the rubber tree, *Hevea brasiliensis*. 
Overview of terpene biogenesis

- Isopentenyl diphosphate (IPP) is the C5 precursor of all isoprenoids. In plants, IPP is formed via two distinct biosynthetic pathways:

  - The **mevalonic acid (MVA) pathway** operates in the cytoplasm and is responsible for the smaller terpenes and the phytosterols.

  - The **methyl erythritol phosphate (MEP) pathway** is responsible for the chloroplast isoprenoids (β-carotene, lutein, prenyl chains of chlorophylls and plastoquinone-9).
The Mevalonic Acid (MVA) pathway

• 3R-Mevalonic acid (MVA) is biosynthesized from three acetates.

Note: 3S-MVA is an unnatural stereoisomer. There is no evidence that it is incorporated into terpenes.

• MVA is converted to *isopentenyl diphosphate* (IPP) which is converted to its isomer, *dimethylallyl diphosphate* (DMAPP).
MVA pathway for isoprenoid biosynthesis with labeling pattern from $[1^{13}C]$glucose metabolized via glycolysis. (Rohmer, Pure Appl Chem 2003)
The Methyl Erythritol Phosphate (MEP) pathway

- Glyceraldehyde-3-phosphate (GAP) and phosphoenolpyruvate (PEP) are formed from glucose.

- GAP condenses with PEP to form MEP. MEP is converted to IPP which forms its isomer DMAPP.
The Methyl Erythritol Phosphate (MEP) pathway

MEP pathway for the biosynthesis of isoprenoids with labeling pattern from [1-$^{13}$C]glucose metabolized via glycolysis. (Rohmer, Pure Appl Chem 2003)
Evolution of the MVA and MEP pathways

• The MVA pathway was originally thought to be the obligatory intermediate for all terpenes. (This is the pathway assumed in pre-2000 literature.)

• The MEP pathway was first found in eubacteria and green algae, and was later shown to operate in the plant’s chloroplast. It is hypothesized that the MEP evolved first, and was incorporated into plants from cyanobacteria.

• Some fungi and yeasts have been shown to use the MVA pathway. Because the plant cytosol uses the MVA pathway, it is believed that the higher evolved organisms (fungi and yeast) may be the source of the plant’s nuclear DNA.

• The co-occurrence of two distinct major metabolic pathways in plant cells is unique for isoprenoid formation in plant cells.

- IPP is “the central intermediate in the biosynthesis of isoprenoids, the most ancient* and diverse class of natural products. Two distinct routes of IPP biosynthesis occur in nature: the MVA pathway and the recently discovered DXP** pathway.”

“The evolutionary history of the enzymes involved in both routes and the phylogenetic distribution of their genes across genomes suggest that:

1. the MVA pathway is germane to archaebacteria,
2. that the DXP pathway is germane to eubacteria, and that eukaryotes have inherited their genes for IPP biosynthesis from prokaryotes.”

* In evolutionary terms, the fats are probably the older group!
** DXP (deoxyxylulose 5-phosphate) pathway = MEP pathway

“The occurrence of genes specific to the DXP pathway is restricted to plastid-bearing eukaryotes, indicating that these genes were acquired from the cyanobacterial ancestor of plastids.

“However, the individual phylogenies of these genes, with only one exception, do not provide evidence for a specific affinity between the plant genes and their cyanobacterial homologues. The results suggest that:

3 lateral gene transfer between eubacteria subsequent to the origin of plastids has played a major role in the evolution of this pathway.”
# The MVA and MEP pathways: taxonomic distribution

<table>
<thead>
<tr>
<th>Organism</th>
<th>Pathways</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacteria</td>
<td>MVA or MEP</td>
</tr>
<tr>
<td>Archaea</td>
<td>MVA</td>
</tr>
<tr>
<td>Green Algae</td>
<td>MEP</td>
</tr>
<tr>
<td>Fungi</td>
<td>MVA</td>
</tr>
<tr>
<td>Plants</td>
<td>MVA and MEP</td>
</tr>
<tr>
<td>Animals</td>
<td>MVA</td>
</tr>
</tbody>
</table>
The MVA and MEP pathways: practical implications

• The mevalonate-independent methylerythritol phosphate (MEP) pathway is present in many bacteria and in the chloroplasts of all phototrophic organisms. It represents an alternative to the well-known MVA pathway, which is present in animals, fungi, plant cytoplasm, archaebacteria, and some eubacteria.

• The MEP pathway in these bacteria represents a novel selective target for antibacterial and antiparasitic drugs.

• The MEP pathway is also present in nonphototrophic eukaryotes, but belonging to phyla related to phototrophic unicellular eukaryotes, such as the parasite responsible for malaria, *Plasmodium falciparum*. This presents a potential target for a new class of antibacterial and antiparasitic drugs.
The MVA and MEP pathways

HMGR, 3-hydroxy-3-methylglutaryl coenzyme A reductase
IDI, isopentenyl diphosphate isomerase

Compartmentalized biosynthesis of IPP and DMAPP via the cytosolic MVA and the plastidic MEP pathways.

DXS, 1-deoxy-d-xylulose-5-phosphate synthase
DXR, 1-deoxy-d-xylulose-5-phosphate reductoisomerase
HDS, hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase
IDS, isopentenyl diphosphate dimethylallyl diphosphate synthase
IDI, isopentenyl diphosphate isomerase

The terpene family is formed by condensation of C5 (IPP) units:

- C10, monoterpenes
- C15, sesquiterpenes
- C20, diterpenes.

DMAPP, C5

DMAPP, C5

geranyl pyrophosphate, C10

farnesyl pyrophosphate, C15

geranylgeranyl pyrophosphate, C20
Terpene chains are produced by condensation of DMAPP with IPP in “head-to-tail” manner. DMAPP is the “starter unit” while IPP is the nucleophile which lengthens the terpene chain.
C30 terpenes are formed by head-to-head dimerization of C15 sesquiterpenes. This leads to the triterpenes, steroids, and carotenes.
Overview of Terpene Biosynthesis in Plants

Cytosol

CoA-SCoA

Glucose

DMAPP, C5

IPP, C5

MVA

Monoterpenes

prenyl side-chain

DMAPP, C5

IPP, C5

geranyl pyrophosphate, C10

farnesyl pyrophosphate, C15

Sesquiterpenes

farnesyl pyrophosphate, C15

Diterpenes

geranylglyceranyl pyrophosphate, C20

Triterpenes & Steroids

squalene, C30

Plastids

GAP

PEP

MEP

DMAPP, C5

IPP, C5

geranyl pyrophosphate, C10

farnesyl pyrophosphate, C15

Sesquiterpenes

polyprenyl side-chain

Monoterpenes

Head-to-head dimerization

Head-to-head dimerization

Carotenoids
Estimates of number of structural groups and compounds known for each of the major types of terpenes. (Devon and Scott, 1972; Dictionary of Terpenoids, 1991)

<table>
<thead>
<tr>
<th>Main terpene group</th>
<th>Number of structural types</th>
<th>Approx. number of known compounds (1991)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monoterpene</td>
<td>8</td>
<td>850</td>
</tr>
<tr>
<td>Sesquiterpene</td>
<td>88</td>
<td>2,800</td>
</tr>
<tr>
<td>Diterpene</td>
<td>53</td>
<td>2,500</td>
</tr>
<tr>
<td>Triterpene</td>
<td>25</td>
<td>1,500</td>
</tr>
<tr>
<td>Phytosteroid</td>
<td>19</td>
<td>800</td>
</tr>
</tbody>
</table>
Phylogenetics and natural products: Evolution of MVA biosynthesis in plants

6. Terpenes: C5 to C20 (Dayrit)

- Iridoids (Labiatae)
- Sesquiterpenes (Myrtaceae)
- Diterpenes (Apocynaceae)
- Sesquiterpene lactones (Compositae)
- Diterpenes (Leguminoseae)
- Steroidal alkaloids (Solanaceae)
- Plumeria spp.
- Jasmine
- Scutellaria
- Daisy
- Tomato
- Lima bean
Common transformations of the terpenes.

1. Sn2-type attack of carbon nucleophile on a carbon electrophile with loss of (-OPP):

\[ \text{Sn2-type attack of carbon nucleophile on a carbon electrophile with loss of (-OPP):} \]

\[ \text{OPP} \]

\[ \text{X} \]

\[ \text{Enz} \]

\[ \text{IPP} \]

\[ \text{DMAPP} \]

2. E2-type elimination reaction with formation of double bond:

\[ \text{E2-type elimination reaction with formation of double bond:} \]

\[ \text{OPP} \]

\[ \text{X} \]

\[ \text{Enz} \]

\[ \text{H} \]

:Base

3. 1,3-Allyl shift of X group (-OPP):

\[ \text{3. 1,3-Allyl shift of X group (-OPP):} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{CH}_3 \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{H} \]

\[ \text{C}_3 \]

\[ \text{O} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{H} \]

\[ \text{C}_3 \]

\[ \text{O} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{H} \]

\[ \text{C}_3 \]

\[ \text{O} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{H} \]

\[ \text{C}_3 \]

\[ \text{O} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{H} \]

\[ \text{C}_3 \]

\[ \text{O} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{H} \]

\[ \text{C}_3 \]

\[ \text{O} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{H} \]

\[ \text{C}_3 \]

\[ \text{O} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{H} \]

\[ \text{C}_3 \]

\[ \text{O} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{H} \]

\[ \text{C}_3 \]

\[ \text{O} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{H} \]

\[ \text{C}_3 \]

\[ \text{O} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{H} \]

\[ \text{C}_3 \]

\[ \text{O} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{H} \]

\[ \text{C}_3 \]

\[ \text{O} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{P} \]

\[ \text{R} \]

\[ \text{R} \]

\[ \text{R} \]
Common transformations of the terpenes.

5. Electrophilic attack on double bond or epoxide to produce carbocation 
\( + (E \text{ is usually } H^-) \):

6. Formation of cyclopropyl group from olefin:
Common transformations of the terpenes.

7. Cationic rearrangements:

a. 1,2-Shift of methyl or hydrogen:

\[ R_1^+ + R_2 \rightleftharpoons R_2^+ + R_1 \]

b. 1,3-Shift of hydrogen:

\[ R_1^+ + R_3 \rightleftharpoons R_3^+ + R_1 \]

c. Acid-catalyzed cationic cyclization:

\[ R \xrightarrow{H^+} \]
Common transformations of the terpenes.

8. Alkylation of alcohol and alkyl groups with isopentenyl group (prenylation); analogous to methylation:
Hemiterpenes, C5

- The true C5 terpenes are known as hemiterpenes. However, there is a large number of C5 compounds that are degradation products of larger compounds and these are not considered as true terpenes.
- **Prenylation** is an alkylation process which adds C5 units to a substrate by attack of a carbon nucleophile or alcohol on the C5 diphosphate \((C_5H_9-\text{OPP})\). The reaction takes place by nucleophilic displacement of diphosphate. If the substrate is a non-terpene compound, a mixed metabolite compound is produced.
Various aspects of C5 terpene chemistry:

A. true hemiterpenes

B. false hemiterpenes

C. examples of prenylation.
Monoterpenes, C10

Monoterpenes are characteristic plant natural products, in particular, in the flowers. These are important materials for perfumes and food flavors. The monoterpenes are classified structurally into the following: acyclic, monocyclic, bicyclic and tricyclic.

Biosynthetically, the monoterpenes can be grouped into the following: acyclic, cyclic, iridoid, and irregular.

Geranyl diphosphate \((C_{10}H_{17}-\text{OPP})\) is the starting point for the monoterpenes.
6. Terpenes: C5 to C20 (Dayrit)

Other common open chain terpenes:
- Citronellal
- Citral A (geranial)
- Geraniol
- Linalool
- Citral B (neral)
- Nerol
- Neryl pyrophosphate
- Linalyl pyrophosphate
- Acyclic monoterpenes

 cis-ocimene
myrcene
Cyclic monoterpenes

- In order to form the cyclic monoterpenes, it is postulated that GPP isomerizes to neryl or linalyl diphosphate. Cyclization produces a 1-methyl 4-isopropyl cyclohexane ring system. This is the basis for limonene-type structures. Ring formation can occur via three cyclization routes which produces bicyclic structures.

- Ring formation produces optically active cyclic monoterpenes. Both (+) and (-) enantiomers have been found for many of the cyclic monoterpenes; these enantiomers can be found in different plants but sometimes both enantiomers can be found in the same plant.
Route to the various mono- and bicyclic monoterpenes.
Cyclization of monoterpenes can occur with opposite conformations giving rise to enantiomers.

D-Limonene:

L-Limonene:
Labeling studies for the conversion of MVA into cyclic monoterpenes.

Labeling convention:
- : full label
- : half-label

Conversion diagram:

MVA → (intermediate) → (intermediate) → sabinene

- Sabinene: 99% full label, 50% half-label

MVA → (intermediate) → (intermediate) → thujone

- Thujone: 90-99% full label, 50% half-label

MVA → (intermediate) → (intermediate) → camphor

- Camphor: 80% full label, 50% half-label

Chemical structures:
- MVA (33 carbon atoms)
- Sabinene (28 carbon atoms)
- Thujone (28 carbon atoms)
- Camphor (28 carbon atoms)
Wagner-Meerwein rearrangement of cationic intermediates in cyclic monoterpenes.
Iridoids

- Iridoids are rearranged monoterpenes which have a characteristic fused 5/6-membered ring structure. The 6-membered ring contains an acetal carbon where one oxygen forms part of the ring as a C-O-C bond, and the other oxygen is usually glycosylated.

- About 600 iridoids are known. However, these are mostly glycosides; only about 100 non-glycosidic iridoids are known.

- Iridoids are generally plant terpenes. The name, however, comes from iridomyrmecin, a compound isolated from ants of the genus Iridomyrmex. Iridoid monoterpenes also are known from other insects, such as aphids.

- Loganin is a typical iridoid. Secologanin is a rearranged iridoid which retains only the 6-membered ring.
Iridoids are rearranged monoterpenes. **Loganin** is a typical example of an iridoid. Opening of the cyclopentane ring yields secologanin. Secologanin is incorporated into numerous indole alkaloids.
Iridoids

- Because the biosynthesis is relatively long and involves steps not commonly seen in other pathways, it would not be expected to have arisen often in the course of evolution.
- The iridoids are produced by plants primarily as a defense against herbivores or against infection by microorganisms. Iridoids are often characterized by a deterrent bitter taste.
- Iridoids are found in many medicinal plants and may be responsible for some of their pharmaceutical activities. Isolated and purified, iridoids exhibit a wide range of bioactivity including cardiovascular, antiheptatoxic, chlorectic, hypoglycemic, anti-inflammatory, antispasmodic, antitumor, antiviral, immunomodulator and purgative activities.
- Iridoids are incorporated into the large family of indole alkaloids.
Irregular monoterpenes

The irregular monoterpenes are a miscellaneous group that include the following types of compounds:

1. compounds that are formed from ring expansion;
2. compounds that are degraded so that the resulting compound has less than 10 carbons; and
3. compounds that are formed via a head-to-head condensation of DMAPP. A well-known group is the pyrethrins which have the characteristic cyclopropyl ring system.
1. Irregular monoterpenes that are formed from ring expansion.
2. Irregular monoterpenes that are degraded so that the resulting compound has less than 10 carbons.
3. Irregular monoterpenes that are formed via head-to-head condensation of DMAPP: chrysanthemic acid. The cyclopropyl group is characteristic of the pyrethroids, which is a well-known insecticide which has low mammalian toxicity.

- Full C-label
- Half C-label

Tetramethrin (an active ingredient in Raid insect spray)
Sesquiterpenes, C15

Sesquiterpenes (*sesqui* = “one and a half”) are derived from C15 farnesyl diphosphate (FPP). The sesquiterpenes comprise a very large group of over 1,000 individual compounds with over 100 skeletal types. The great variety of structures arises from the following types of transformation:

1. There are 3 double bonds in FPP; two of these double bonds can isomerize into *cis* ↔ *trans* configurations. This gives 4 double-bond geometric isomers: all-*trans*; 2-*cis*,6-*trans*; 2-*trans*,6-*cis*; and 2-*cis*,6-*cis* isomers. The most important geometric isomers are the all-*trans* and 2-*cis*,6-*trans* isomers. Each double bond isomer gives rise to a different branch of the sesquiterpene family.
Many sesquiterpenes, however, are derived from either the all-trans isomer, 1, or the 2-cis, 6-trans isomer, 2.
2. When FPP cycles, the position of initial cyclization is the second major source of variation. There are several ways in which FPP can fold leading to a number of cyclization modes. Cyclization can involve either nucleophilic displacement of \( \text{OPP} \) by any of the \( \pi \)-orbitals of the double bonds, or nucleophilic attack of the double bond (usually the terminal double bond) on a proton or other electrophile.

Two main groups arise from cyclization of:

- all-\( \text{trans} \) FPP
- 2-\( \text{cis} \), 6-\( \text{trans} \) FPP
Skeletal types obtained from all-trans FPP.
3. Farnesyl diphosphate is achiral. However, as in the case of geranyl diphosphate, it can fold up into two conformational forms which give rise to enantiomers; the conformational isomers are *pro-chiral*. Further modification of these enantiomers leads to diastereomers which are now chemically distinct compounds.
4. The remaining double bonds can react forming more C-C bonds leading to bicyclic systems. The stereochemistry of the second cyclization and the conformation of folding gives rise to further types of isomers. From the all-trans FPP arises the germacrane subgroup.

A. Simple germacrane metabolites:

B. Overview of three secondary cyclization modes, $X^-$ nucleophile is usually OH$^-$. 

Mode A: 

Mode B: 

Mode C:
Metabolites from all-trans FPP: germacrane skeleton and cyclization mode A: overview.
Metabolites from all-trans FPP: germacrane skeleton and cyclization mode A:

i. chair conformation, base attack from α-face.

ii. chair conformation, base attack from β-face.

iii. boat conformation, base attack from β-face.
Metabolites from all-\textit{trans} FPP: germacrane skeleton and cyclization mode A:

i. chair conformation, base attack from $\alpha$-face.

ii. \textbf{chair conformation, base attack from $\beta$-face.}

iii. boat conformation, base attack from $\beta$-face.
iii. boat conformation, base attack from β-face:

Metabolites from all-*trans* FPP: germacrane skeleton and cyclization mode A:
i. chair conformation, base attack from α-face.
ii. chair conformation, base attack from β-face.
iii. **boat conformation, base attack from β-face.**
Metabolites from all-trans FPP: germacrane skeleton and cyclization mode B.

patchoulenone

bulnesol
Skeletal types obtained from 2-cis, 6-trans FPP (1)
Skeletal types obtained from 2-cis, 6-trans FPP (2)

Part of a biogenetic map for sesquiterpenes showing all structural types found in the oils from *Toona ciliata*, *Cedrela odorata*, and *C. fissilis*. (modified from: http://www.scielo.br/img/fbpe/jbchs/v11n6/3595f3.gif)
Sesquiterpenes, C15

5. Various skeletal rearrangements may occur through a number of mechanisms, such as 1,2- and 1,3-H shifts, 1,2-methyl shifts, double-bond migration (this generally occurs via H migration), and Wagner-Meerwein type rearrangement.

6. Other modifications such as oxidation, reduction, etc.

The sesquiterpenes serve as plant defense compounds (e.g., polygodial, β-eudesmol), sensory attractants or fragrance odors (e.g., ylanganes, patchoulenone). The structures displayed by these compounds attest to the rich chemistry that is found in plants.
The major skeletal types which contain the \textit{α-methylene-γ-lactone} functionality belong to the germacrane group.
Sesquiterpene $\alpha$-methylene-$\gamma$-lactones

- The $\alpha$-methylene-$\gamma$-lactone group is derived by oxidation and cyclization of the isopropyl side chain of the cyclized sesquiterpene.

- The $\alpha$-methylene-$\gamma$-lactone functionality is characteristic of sesquiterpenes, and is found most widely in the Compositae family.
• This distinctive structural feature is accompanied by the observation that many of these compounds show antimutagenic properties which are attributed to the electrophilic exocyclic methylene group which can react rapidly with nucleophilic moieties, such as the nucleophilic 8-carbon of adenine and guanine in DNA.

Reaction scheme for attack by nucleophile, X:
Examples of sesquiterpene $\alpha$-methylene-$\gamma$-lactones.

- Tagitenin A (germacrenolide)
- Hyporadiolide (guaianolide)
- Miloanokrypten (guaianolide)
- Helenalin

Blumealactones from *Blumea balsamifera* (sambong) (Fujimoto, et al., *Phytochem.*, **27**, 1109 (1988)).
Terminal cyclohexyl ring formation: Abscisic acid

- A terminal cyclohexyl group is readily formed from 2-trans, 6-trans farnesyl diphosphate by electrophilic attack. This relatively simple cyclization mode is surprisingly not common among the sesquiterpenes. The best representative of this group is abscisic acid.

- **Abscisic acid** ("abscission" ≡ shedding of leaves, fruits or flowers) is a very important plant growth regulator. In particular, abscisic acid (commonly called ABA), as its name suggests, is the plant hormone responsible for dormancy of leaves and the abscission of leaves, flowers and fruit (the natural process of removal, cutting, or falling off). Thus, ABA plays an important role in normal plant development. In tissue cultures, ABA has also been shown to inhibit plant cell elongation.
Terminal cyclohexyl ring formation leads to the biosynthesis of the important plant growth regulator, abscisic acid.
Sesquiterpenes from cascading “linear” cyclization: Polygodial

- 2-Trans, 6-trans farnesyl diphosphate can cyclize in a cascading “linear” conformation to form a trans-decalin structure. This process is initiated by attack of an electrophile at the terminal double bond, formation of a cyclohexyl ring followed by a second cyclization to form the fused trans-decalin structure.
- This mode of cyclization is commonly observed in the longer diterpenes and triterpenes, but is unusual in sesquiterpenes.
Appropriate folding of all-\textit{trans} farnesyl diphosphate and electrophilic attack at the terminal double bond with sequential cyclization leads to a \textit{trans}-decalin structure. The -OPP\textsuperscript{-} group is not displaced during the initial cyclization. This mode of cyclization is very common for diterpenes and triterpenes, and unusual in sesquiterpenes. Polygodial is an antipest compound produced by plants.
Sesquiterpenes from 2-cis, 6-cis-farnesyl diphosphate:
Gossypol

A few biologically important sesquiterpenes are formed from 2-cis, 6-cis-farnesyl diphosphate. The best known metabolite from this group is gossypol. Gossypol is an unusual sesquiterpene since it is has a naphthalene structure and is dimerized. Although gossypol can be easily mistaken for a polyketide, the isopropyl group hints at its terpenoid origins. The final proof of biogenetic origin comes from labeling studies which are consistent with its being a sesquiterpene metabolite.
Gossypol is formed from 2-cis, 6-cis-farnesyl diphosphate. It is a dimeric naphthalene. Gossypol has attracted interest because it is an insecticidal defense compound found in the seeds of the cotton plant.

\[\begin{align*}
  &\text{Gossypol} \\
  &\text{2-cis, 6-cis-farnesyl diphosphate}
\end{align*}\]

- \(\bullet\) = \(^{14}\)C label
- \(\circ\) = partially labeled
Insect development is characterized by discrete stages in its life cycle in going from larva to adult. Two types of hormones – juvenile hormones (JH) and moulting hormones (MH) -- initiate these changes. Juvenile hormones are required at the initial metamorphosis from the 1\textsuperscript{st} to 2\textsuperscript{nd} stage larva, while molting hormones are required at all stages of development:

\[
\text{MH/JH} \quad \rightarrow \quad \text{MH} \quad \rightarrow \quad \text{MH}
\]

1\textsuperscript{st} stage Larva \quad \rightarrow \quad 2\textsuperscript{nd} stage Larva \quad \rightarrow \quad \text{Pupa} \quad \rightarrow \quad \text{Adult}

Many juvenile hormones use propionyl CoA as the starting unit to form homo-mevalonic acid. Juvenile hormones are linear sesquiterpenes with various oxidized groups (epoxides, alcohols). Moulting hormones, on the other hand, are steroidal compounds.
Juvenile hormones are linear sesquiterpenes which use a propionyl CoA as the starting unit to make homo mevalonic acid.

Starter unit: propionyl CoA:

- Neotenin is a juvenile hormone isolated from thousands of male butterflies of the species *Hyalophora cecropia* L.
Diterpenes, C20

Addition of another isopentenyl diphosphate (C5) group to farnesyl diphosphate (C15) forms geranylgeranyl diphosphate (GGPP, C20). GGPP is the starting point for the biogenesis of the diterpenes. As in the case of the sesquiterpenes, there is a great variety of structures that are formed due to five important structural features of diterpenes:

1. GGPP has four double bonds. Three of the double bonds can take the cis- or trans- configuration. This gives seven possible double-bond geometric isomers.
   - The largest group of diterpenes is formed from the all-trans GGPP.
   - A number of important diterpenes arise from the 2-trans,6-cis,10-trans isomer.
The majority of the diterpenes are cyclic compounds. However, some unusual open-chain di-, tri- and tetraene diterpenes (along with sesquiterpenes) have been isolated from the skin glands of alligators and crocodiles. These terpenes are believed to act as pheromones. (The most commonly-occurring open-chain terpene on the skin of many animals, including humans, is squalene.) (Schultz, Krückert and Weldon, J. Nat Prod., 2003, 66, 34-38)
Diterpenes, C20

2. Initial cyclization of GGPP can occur at different sites. As in the case of FPP, the position of initial cyclization contributes to structural variety. This includes the important consideration of whether the cyclization involves the displacement of the -OPP group or not.

• Cyclization in the all-\textit{trans} mode does not displace -OPP, while the other modes cause displacement of the -OPP group.

3. Further cyclization steps may occur using the remaining double bonds to form additional smaller rings. The stereochemistry of these cyclization steps depends on the double-bond configurations and the conformation of the ring.
GGPP cyclizes in a variety of ways. Theoretically, there are seven double-bond geometric isomers possible from geranylgeranyl diphosphate. The most important diterpenes are formed from the all-trans GGPP and 2-trans, 6-cis, 10-trans isomer.
Diterpenes

4. **Skeletal modifications** can occur via migration of hydrogen and/or methyl, Wagner-Meerwein-type rearrangements, and others.

5. **Other modifications** may occur, such as oxidation, reduction of double bond (+2[H]), etc.

There are a number of well-known compounds which belong to the diterpene group, in particular, the *cembranes*, *pimaranes* and the *gibberelanes*. The cembrane structure is a recurring theme among marine natural products. The pimaranes are major constituents of pine oleoresin, which is heavily used in paper sizing and other coating applications. The gibberellins are important plant growth hormones. Taxol, the anticancer drug, is also a diterpene.
Diterpenes from all-trans GGPP

- The largest group of diterpenes is produced from all-trans GGPP which lead to the labdanes. The biosynthesis starts from two conformations of chair-like folding of all-trans GGPP: one conformation leads to the 10-α series while the alternative conformation leads to the 10-β series. This is an example of a very important phenomenon in generation of diversity in natural products where the same starting compound produces stereoisomeric products—in this case, diastereomers.

- Electrophilic attack at the terminal double bond accompanied by two rapid ring-forming steps produces the two trans-decalin stereoisomers with a stable tertiary cation intermediate. It should be noted that initial cyclization retains the -OPP group which is used in subsequent transformations.
The largest group of diterpenes is produced from all-\emph{trans} GGPP which leads to the labdanes. The biosynthesis starts from two conformations of chair-like folding of all-\emph{trans} GGPP leading to the 10-\(\alpha\)/10-\(\beta\) series.
Diterpenes from all-trans GGPP

The initially formed decalin cation subsequently undergoes various transformations leading to sub-groups, as illustrated by the following reactions. Examples from both the 10-α and 10-β series are known in nature, but both are rarely found in the same plant.

1. **Loss of H⁺**: Loss of H⁺ limits modifications of the ring to the decalin system. This produces an exocyclic methylene group at the 8-position of the decalin. Secondary modifications may take place at the side chain. This group is exemplified by the labdanes.
Loss of H⁺ from the intermediate decalin produces an exocyclic methylene group at the 8-position. Secondary modifications may take place at the side chain. This leads to the labdanes.
Diterpenes from all-trans GGPP

2. Attack of nucleophile (e.g., OH\textsuperscript{-}) at intermediate cation on 8-position: Attack of hydroxide (OH\textsuperscript{-}) at the cationic carbon at the 8-position is controlled by steric considerations. Because the 10-\(\alpha\) methyl blocks approach from the bottom side, the hydroxide has the 8-\(\beta\) configuration. Similarly, the 10-\(\beta\) methyl series gives rise to the 8-\(\alpha\) hydroxide configuration.
Attack of hydroxide (OH\textsuperscript{-}) at the cationic carbon at the 8-position is controlled by steric considerations. The 10-\(\alpha\) methyl blocks nucleophilic attack from the bottom, while the 10-\(\beta\) methyl blocks attack from the top.
3. Tricyclic diterpenes: The 8-methylene exocyclic group participates in a nucleophilic attack on the $-\text{O}^+\text{P}$ group on the side chain and forms the pimaryl tricyclic system. Two stereoisomeric groups of pimaranes are formed from the 10-α and 10-β methyl series.
The 8-methylene exocyclic group participates in a nucleophilic attack on the \(-\text{OPP}\) group on the side chain and form the pimaryl tricyclic system. This figure shows the pimarane 10-\(\alpha\) and 10-\(\beta\) series.
The 8-methylene exocyclic group participates in a nucleophilic attack on the -OPP group on the side chain and form the pimaryl tricyclic system.
4. Tetracyclic diterpenes: Further cyclization of the tricyclic cationic intermediate leads to a tetracyclic cationic intermediate known as the kaurenes. Two stereoisomeric groups are formed. Many of the kaurenes display remarkable complexity of structure.
Further cyclization of the tricyclic cationic intermediate leads to a tetracyclic cationic intermediate known as the kaurenes. This figure features the kaurane 10-α series.
Further cyclization of the tricyclic cationic intermediate leads to a tetracyclic cationic intermediate known as the kaurenes. This figure features the kaurane 10-β series.
5. Kaurene B-ring contraction: the gibberellins: From the 10α–kaurene series comes the important group of plant hormones, the gibberellins. The gibberellins are synthesized in the protoplasm of plants and increase the rate and amount of growth. There are over 66 compounds which have been isolated which belong to this structural group. Interestingly, gibberellic acid was first isolated in a fungus, *Gibberella*.
Contraction of the kaurene B-ring leads to the gibberellins. The gibberellins are plant growth hormones which increase the rate and amount of growth.
Isomerization of the 6-double bond of all-trans GGPP to the cis configuration allows nucleophilic attack of the terminal double bond on the 1-position accompanied by displacement of -OPP and formation of a 14-membered cembrane ring.
The name cembrane is taken from the simplest member of this group, cembrene, which was isolated form the oleoresin of *Pinus sibirica*.

The most interesting members of the cembrane group are found in marine soft corals. The 14-carbon ring marine cembranes are in many ways structurally analogous to the 10-carbon ring sesquiterpenes. While the sesquiterpene germacranes form $\alpha$-methylene-$\gamma$-lactones, the diterpene cembranes form $\alpha$-methylene-$\delta$-lactones. Like the sesquiterpene lactones, the cembrane lactones also exhibit powerful cytotoxic and anti-tumor properties.
Isomerization of the 6-double bond of all-trans GGPP to the cis configuration allows nucleophilic attack of the terminal double bond on the 1-position accompanied by displacement of –OPP– and formation of the cembrane ring.

Cembrene itself was isolated from the oleoresin of Pinus sibirica. 

α-Methylene-δ-lactones in marine natural products:

Crassin acetate from Psuedoplaxaura porosa

Sinulariolide from Sinularia flexibilis
Diterpenes from all-\textit{trans} GGPP with loss of OPP: Taxanes

The biosynthetic pathway to the tricyclic taxoid skeleton involves an alternative cyclization attack of C10 on C15, accompanied by attack of C14 on C1 with displacement of \(-\text{OPP}\). This forms a 6-membered A ring, with an 8-membered B-ring.
The taxoid skeleton is formed from all-\textit{trans} GGPP. The initial step involves attack of C14 on C1 with displacement of -OPP and attack of C10 on the cationic center which forms on C15.
Taxol is a well-known diterpenoid from the Pacific yew tree, *Taxus brevefolia*. Over 350 compounds belonging to the taxoid group have been isolated. In a short span of seven years from 1992-1999, over 250 taxoids were isolated and characterized. (Baloglu and Kingston, *J. Nat. Prod.* **1999**, *62*, 1448-1472.)
The nonmevalonate pathway supports both monoterpenes and sesquiterpene formation in snapdragon flowers

Natalia Dudareva**, Susanna Andersson*, Irina Orlova*, Nathalie Gatto†, Michael Reichelt‡, David Rhodes*, Wilhelm Boland‡, and Jonathan Gershenzon‡

Potential labeling patterns of monoterpenes and the sesquiterpene nerolidol by $^{[2\text{H}_2]}$-DOX (A) and $^{[2\text{H}_2]}$-MVL (B).
Emission and *in vivo* labeling kinetics of nerolidol during feeding of snapdragon flowers with $[^2\text{H}_2]$-MVL. Filled circles represent emission from $[^2\text{H}_2]$-MVL-fed flowers, and open circles represent control flowers fed with sucrose solution. Open triangles represent percent of unlabeled compounds, and filled squares and triangles represent percent of nerolidol labeled by +4 amu; open squares represent percent of nerolidol labeled by +6 amu.

In the case of snapdragon petals, only one of the two pathways is operating in the formation of volatiles isoprenoids. The MEP pathway, localized in the plastids, provides IPP and DMAPP precursors for both monoterpene biosynthesis (in plastids) and sesquiterpene biosynthesis (in the cytosol) and determines their rhythmic emission.
Summary

The following biosynthetic patterns among terpenes can be observed:

1. The terpenes are readily recognizable from the characteristic presence of 1,5-methyl substituents and the carbon number which is a multiple of 5.

2. The key intermediates, **geranyl diphosphate**, **farnesyl diphosphate** and **geranylgeranyl diphosphate**, have a -OPP group at the C1 position and 1,5-diene groups.

3. **Monoterpenes**: There are three main types: linear, cyclic and the iridoids.

4. **Sesquiterpenes**: This a very diverse group of over 100 skeletal types. The initial source of diversity is the variation of the cyclization step.
Summary

5. **Diterpenes**: Diversity starts from the initial cyclization step from the four double bonds in geranylgeranyl diphosphate. The most prominent subgroups are the labdanes which arise from cyclization of the all-trans GGPP, the cembranes which come from cyclization of 2-trans,6-cis,10-trans GGPP, and the taxanes which arise from cyclization of all-trans GGPP with loss of -OPP.

Terpene chemistry is an exquisite example of diversity, stereochemistry and biochemical control in natural products. Despite the diversity of structure, however, the majority of chemical reactions are actually limited to a few types: Sn2 displacement of -OPP by an olefinic group; E2 elimination to yield an olefin; 1,3-allyl shift of -OPP; epoxidation; and protonation of olefins followed by cationic rearrangement.
Summary

6. **MVA and MEP Pathways:** The co-occurrence of two completely distinct pathways for isoprenoid formation in plant cells is remarkable because a similar situation does not hold for any other major metabolic route. The plastidial pathway probably arose from genes contained in a cyanobacterium-like symbiont that served as the progenitor of modern chloroplasts. However, this scenario still does not explain the persistence of both pathways in contemporary plants. The answer may lie in the enormous variety of isoprenoids formed by plants, which could require two separate pathways composed of completely different enzymes and different intermediates to facilitate separate regulation. Further study of when and where the two pathways are active in plants should shed further light on questions regarding their evolutionary origin and maintenance. (Dudareva et al. PNAS 2005)