Triterpene Biosynthesis in Plants

Ramesha Thimmappa, Katrin Geisler, Thomas Louveau, Paul O’Maille, and Anne Osbourn

1Department of Metabolic Biology, John Innes Centre, Norwich Research Park, Norwich NR4 7UH, United Kingdom; email: anne.osbourn@jic.ac.uk
2Food and Health Program, Institute of Food Research, Norwich Research Park, Norwich NR4 7UH, United Kingdom

Keywords
oxidosqualene cyclases, metabolic gene clusters, specialized metabolites, sterols, synthetic biology

Abstract
The triterpenes are one of the most numerous and diverse groups of plant natural products. They are complex molecules that are, for the most part, beyond the reach of chemical synthesis. Simple triterpenes are components of surface waxes and specialized membranes and may potentially act as signaling molecules, whereas complex glycosylated triterpenes (saponins) provide protection against pathogens and pests. Simple and conjugated triterpenes have a wide range of applications in the food, health, and industrial biotechnology sectors. Here, we review recent developments in the field of triterpene biosynthesis, give an overview of the genes and enzymes that have been identified to date, and discuss strategies for discovering new triterpene biosynthetic pathways.
INTRODUCTION

Sterols and triterpenes are isoprenoids that are synthesized via the mevalonate pathway (21). The last common intermediate for their two pathways is 2,3-oxidosqualene. Sterols are important structural components of membranes and also have roles in signaling (as steroidal hormones). In contrast, triterpenes are not regarded as essential for normal growth and development, and although they do exist in plants in simple unmodified form, they often accumulate as conjugates with carbohydrates and other macromolecules, most notably as triterpene glycosides. Triterpene glycosides have important ecological and agronomic functions, contributing to pest and pathogen resistance and to food quality in crop plants. They also have a wide range of commercial applications in the food, cosmetics, pharmaceutical, and industrial biotechnology sectors (6, 88, 99, 119, 147).

The cyclization of squalene (in bacteria) or 2,3-oxidosqualene (in fungi, animals, and plants) to sterol or triterpene products is one of the most complex enzymatic reactions known in terpene metabolism (1, 104, 153). These reactions have intrigued organic chemists and biochemists for the past half century (26). For the purposes of this review, we make a distinction between sterols and triterpenes based on the way in which these molecules are synthesized. In sterol biosynthesis, 2,3-oxidosqualene is cyclized to the sterols lanosterol (in fungi and animals) or cycloartenol (in plants) via the chair-boat-chair (CBC) conformation. In triterpene biosynthesis, in contrast, this substrate is folded into a different conformation—the chair-chair-chair conformation (CCC)—prior to cyclization into a huge array of triterpenes of diverse skeletal types, of which just one (β-amyrin) is shown as an example in Figure 1.

The triterpenes are one of the largest classes of plant natural products, with more than 20,000 different triterpenes reported to date (47). The vast majority of triterpene diversity occurs in the plant kingdom, although other organisms also produce triterpenes. Examples include the synthesis of the simple triterpene hopene from squalene in bacteria (101) and the production of defense-related triterpene glycosides by sea cucumbers (146). Some other plant-derived specialized metabolites are synthesized via the CBC conformation, such as cucurbitacin (associated...
Figure 1
The biosynthetic route to sterols and triterpenes. Sterols and triterpenes are synthesized via the mevalonic acid (MVA) pathway. The enzymes that catalyze the various steps are indicated in boxes. Enzyme abbreviations: FPS, farnesyl pyrophosphate synthase; SQS, squalene synthase; SOE, squalene monooxygenase or epoxidase; SHC, squalene-hopene cyclase; LAS, lanosterol synthase; CAS, cycloartenol synthase; CPQ, cucurbitadienol synthase; BAS, β-amyrin synthase. Other abbreviations: CBC, chair-boat-chair; CCC, chair-chair-chair.
with the bitter taste of many members of the Cucurbitaceae family), which are synthesized from
curcurbitadienol (130) (Figure 1), and **tomato steroidal glycoalkaloids**, which are synthesized via
cholesterol (54).

More than 100 different triterpene scaffolds are currently known in plants (163). Triterpene cy-
clization can thus lead to a wide array of different triterpene structures, all derived from the simple
and ubiquitous linear isoprenoid substrate 2,3-oxidosqualene. These triterpene scaffolds can then
provide the foundation for further modification by triterpene-modifying (or tailoring) enzymes
(e.g., cytochrome P450s, sugar transferases, and acyltransferases), thereby leading to enormous
structural diversity. Nature’s triterpene reservoir remains largely undiscovered, despite the con-
siderable commercial interest in these compounds for a range of applications (6, 88, 99, 119).

Major advances in our understanding of the genes, enzymes, and pathways required to synthesize
these molecules are now opening up unprecedented opportunities for triterpene metabolic engi-
neering and for the discovery of new pathways and chemistries, facilitated by the recent discovery
that the genes for triterpene pathways are—in at least some cases—organized in biosynthetic gene
clusters in plant genomes (28, 29, 63, 69, 107).

**TRITERPENE BIOSYNTHESIS**

**Cyclization of 2,3-Oxidosqualene: One Substrate, an Array of Products**

Cyclization of 2,3-oxidosqualene is catalyzed by enzymes known as oxidosqualene cyclases (OSCs),
which generate either sterol or triterpene scaffolds in a process involving (a) substrate binding and
preorganization (folding), (b) initiation of the reaction by protonation of the epoxide, (c) cyclization
and rearrangement of carbocation species, and (d) termination by deprotonation or water capture
to yield a final terpene product (Figure 2). Although variation in carbocation cyclization and
rearrangement steps contributes substantially to scaffold diversity, the initial substrate folding
step is critical, because this predisposes the substrate to follow a particular cyclization pathway.
For example, the CBC conformation organizes cyclization to form the protosteryl cation, which
then gives rise to sterols, whereas the CCC conformation directs cyclization into the dammarenyl
cation, which then gives rise to a host of diverse triterpene scaffolds. In the synthesis of triterpenes
in bacteria, squalene is cyclized to pentacyclic hopene by squalene-hopene cyclases (SHCs) (1).

Following substrate folding, SHCs and OSCs initiate the cyclization reaction by protonation of
the terminal double bond of squalene and the epoxide of 2,3-oxidosqualene, respectively. This protonation step defines SHC and OSC enzymes as class II terpene synthases. By contrast, class I terpene synthases (such as mono-, sesqui-, and some diterpene synthases) initiate cyclization through ionization of pyrophosphate with the assistance of Mg$^{2+}$ cofactors (98).

**Characterized Oxidosqualene Cyclases**

More than 80 OSCs have now been functionally characterized from plants, mostly by heterologous expression of cDNAs in appropriate yeast strains (Table 1). Approximately one-third of these are sterol synthases, a group that includes cycloartenol synthases as well as several lanosterol synthases. The function of lanosterol synthases in plants is unclear, but these enzymes appear to be involved to a small extent in the synthesis of phytosterols and potentially also steroid-derived metabolites (68, 97, 141). Some plant OSCs synthesize molecules other than cholesterol and lanosterol via the CBC fold; examples include cucurbitadienol synthase (CPX) from *Cucurbita pepo* (130) and parkeol synthase (OsPS) from rice (*Oryza sativa*) (56) (Table 1). The remaining OSCs synthesize triterpenes via the substrate CCC conformation (Figures 1 and 2). Collectively, plant OSCs are able to make a diverse array of triterpene scaffolds. Some make common scaffolds such as β-amyrin and lupeol, whereas others generate a host of other single or multiple 2,3-oxidosqualene cyclization products (Table 1).

OSC are encoded by multigene families in plants, so a single plant species is likely to be able to make multiple triterpene scaffolds. For example, the *Arabidopsis thaliana* genome contains 13 OSC genes, including genes for cycloartenol synthase (AtCAS1); lanosterol synthase (AtLSS1); a β-amyrin synthase (LUP4); several mixed-function OSCs that make β-amyrin, lupeol, and a range of other products; and other OSCs that produce “speciality” triterpenes such as manneral (MRN1), thalianol (THAS1), and arabidiol (PEN1) (22, 25, 27, 52, 66–68, 72, 73, 77, 87, 126, 132, 135, 141) (Table 1). These 13 OSCs have different expression patterns and make different major products (Figure 3, Table 1), suggesting that they have specialized functions. BARS1 from *A. thaliana* makes baroul as its predominant cyclization product (~90% abundance) but also makes 22 additional products (all at <2% abundance) (77) (Table 1). It is remarkable that BARS1 is able to make a total of 23 diverse cyclic products, including monocycles (6), bicycles (6/6), tricycles (6/6/6), tetracycles (6/6/6/5 and 6/6/6/6), and pentacycles (6/6/6/6/5 and 6/6/6/6/6). The BARS1 active site is thus able to deprotonate numerous sites from the A to the E ring, possibly at 14 distinct sites (77). The formation of multiple products by BARS1 may reflect the failure of this enzyme to precisely control cyclization toward baroul. The function of OSCs may be to prevent alternative cyclization paths rather than to stabilize particular intermediates that are being directed toward predetermined products—a concept known as negative catalysis (111). Thus, mechanistic diversity may be the default for triterpene cyclization, and product accumulation may result from the exclusion of alternative pathways (77).

A friedelin synthase (FRS) has recently been cloned from *Kalanchoe daigremontiana* (150). FRS is able to span the maximum range of possible skeletal rearrangement from the dammarenyl cyclization pathway (Figure 2). Movement of the positive charge from the C-20 to the C-2 position involves the maximum possible number of 1,2 shifts (10 in total). When the cation reaches the C-2 position, it is attacked by the 3β-OH group to give a ketone group at C-3. Friedelin is therefore one of the most highly rearranged triterpenes known in plants, and unlike most other pentacyclic triterpenes, it lacks a double bond.

Most triterpene synthases follow the standard route of C-C ring-forming and skeletal rearrangement reactions, in the process establishing several stereocenters and forming triterpene alcohols by introduction of a 3β-OH group and a double bond. However, FRS is not the only
Table 1  Plant oxidosqualene cyclases (OSCs)

<table>
<thead>
<tr>
<th>OSC</th>
<th>Species</th>
<th>GenBank ID/gene number</th>
<th>Product</th>
<th>Fold</th>
<th>Ring system</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSS1</td>
<td>Arabidopsis thaliana</td>
<td>At3g45130</td>
<td>Lanosterol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>68</td>
</tr>
<tr>
<td>LAS</td>
<td>Lotus japonicus</td>
<td>AB244671</td>
<td>Lanosterol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>118</td>
</tr>
<tr>
<td>PNZ1</td>
<td>Panax ginseng</td>
<td>AB009031</td>
<td>Lanosterol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>141</td>
</tr>
<tr>
<td>CAS1</td>
<td>Abies magnifica</td>
<td>AF216755</td>
<td>Cycloartenol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>—</td>
</tr>
<tr>
<td>ACX</td>
<td>Adiantum capillus-ceneris</td>
<td>AB368375</td>
<td>Cycloartenol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>137</td>
</tr>
<tr>
<td>CAS1</td>
<td>Arabidopsis thaliana</td>
<td>At2g07050</td>
<td>Cycloartenol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>22</td>
</tr>
<tr>
<td>CS1</td>
<td>Avena strigosa</td>
<td>AJ311790</td>
<td>Cycloartenol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>40</td>
</tr>
<tr>
<td>BPX1</td>
<td>Betula platyphylla</td>
<td>AB055509</td>
<td>Cycloartenol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>167</td>
</tr>
<tr>
<td>BPX2</td>
<td>Betula platyphylla</td>
<td>AB055510</td>
<td>Cycloartenol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>167</td>
</tr>
<tr>
<td>CAS1</td>
<td>Chlamydomonas reinhardtii</td>
<td>EDP09612</td>
<td>Cycloartenol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>80</td>
</tr>
<tr>
<td>CsOSC1</td>
<td>Costus speciosus</td>
<td>AB058507</td>
<td>Cycloartenol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>61</td>
</tr>
<tr>
<td>CPX</td>
<td>Cucurbita pepo</td>
<td>AB116237</td>
<td>Cycloartenol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>130</td>
</tr>
<tr>
<td>CAS1</td>
<td>Dioscorea zingiberensis</td>
<td>AM697885</td>
<td>Cycloartenol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>—</td>
</tr>
<tr>
<td>CAS1</td>
<td>Glycyrrhiza glabra</td>
<td>AB025968</td>
<td>Cycloartenol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>45</td>
</tr>
<tr>
<td>KdCAS</td>
<td>Kalanchoe daigremontiana</td>
<td>HM623872</td>
<td>Cycloartenol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>150</td>
</tr>
<tr>
<td>CAS</td>
<td>Kandelia candel</td>
<td>AB292609</td>
<td>Cycloartenol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>8</td>
</tr>
<tr>
<td>OSC5</td>
<td>Lotus japonicus</td>
<td>AB181246</td>
<td>Cycloartenol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>120</td>
</tr>
<tr>
<td>CAS1</td>
<td>Luffia cylindrica</td>
<td>AB033334</td>
<td>Cycloartenol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>42</td>
</tr>
<tr>
<td>OSC2</td>
<td>Oryza sativa</td>
<td>AK121211</td>
<td>Cycloartenol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>56</td>
</tr>
<tr>
<td>PNX</td>
<td>Panax ginseng</td>
<td>AB009029</td>
<td>Cycloartenol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>71</td>
</tr>
<tr>
<td>CASPEA</td>
<td>Pison sativum</td>
<td>D96919</td>
<td>Cycloartenol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>86</td>
</tr>
<tr>
<td>CAS</td>
<td>Pohyiodoidea nipponica</td>
<td>AB530328</td>
<td>Cycloartenol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>—</td>
</tr>
<tr>
<td>CAS</td>
<td>Rhizophora stylosa</td>
<td>AB292608</td>
<td>Cycloartenol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>8</td>
</tr>
<tr>
<td>ReCAS</td>
<td>Ricinus communis</td>
<td>DQ268870</td>
<td>Cycloartenol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>35</td>
</tr>
<tr>
<td>CPQ</td>
<td>Cucurbita pepo</td>
<td>AB116238</td>
<td>Cucurbitadienol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>130</td>
</tr>
<tr>
<td>OsPS</td>
<td>Oryza sativa</td>
<td>AK066327</td>
<td>Parkeol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>56</td>
</tr>
<tr>
<td>OEA</td>
<td>Olea europea</td>
<td>AB291240</td>
<td>α-Amyrin</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>112</td>
</tr>
<tr>
<td>LUP4</td>
<td>Arabidopsis thaliana</td>
<td>At1g78950</td>
<td>β-Amyrin</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>132</td>
</tr>
<tr>
<td>bAS</td>
<td>Artemisia annua</td>
<td>EU330197</td>
<td>β-Amyrin</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>62</td>
</tr>
<tr>
<td>OX1</td>
<td>Aster selvifolius</td>
<td>AY836006</td>
<td>β-Amyrin</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>14</td>
</tr>
<tr>
<td>bAS1</td>
<td>Avena strigosa</td>
<td>AJ111789</td>
<td>β-Amyrinb</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>40</td>
</tr>
<tr>
<td>BPY</td>
<td>Betula platyphylla</td>
<td>AB055512</td>
<td>β-Amyrin</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>167</td>
</tr>
<tr>
<td>bAS</td>
<td>Bruguiera gymnorrhiza</td>
<td>AB289585</td>
<td>β-Amyrin</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>8</td>
</tr>
<tr>
<td>AS</td>
<td>Euphorbia turcullii</td>
<td>AB206469</td>
<td>β-Amyrin</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>60</td>
</tr>
<tr>
<td>bAS1</td>
<td>Glycyrrhiza glabra</td>
<td>AB037203</td>
<td>β-Amyrin</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>44</td>
</tr>
<tr>
<td>AMY1</td>
<td>Lotus japonicus</td>
<td>AB181244</td>
<td>β-Amyrin</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>120</td>
</tr>
<tr>
<td>AMY1</td>
<td>Medicago truncatula</td>
<td>AJ430607</td>
<td>β-Amyrinb</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>57</td>
</tr>
<tr>
<td>bAS1</td>
<td>Nigella sativa</td>
<td>FJ013228</td>
<td>β-Amyrins</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>124</td>
</tr>
<tr>
<td>PNY1</td>
<td>Panax ginseng</td>
<td>AB009030</td>
<td>β-Amyrin</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>71</td>
</tr>
<tr>
<td>PNY2</td>
<td>Panax ginseng</td>
<td>AB014057</td>
<td>β-Amyrin</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>71</td>
</tr>
</tbody>
</table>

*(Continued)*
Table 1 (Continued)

<table>
<thead>
<tr>
<th>OSC</th>
<th>Species</th>
<th>GenBank ID/gene number</th>
<th>Product</th>
<th>Fold</th>
<th>Ring system</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSY</td>
<td>Pisum sativum</td>
<td>AB034802</td>
<td>β-Amyrin</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>86</td>
</tr>
<tr>
<td>bAS</td>
<td>Polygala tenuifolia</td>
<td>EF107623</td>
<td>β-Amyrin</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>—</td>
</tr>
<tr>
<td>TTS1</td>
<td>Solanum lycopersicum</td>
<td>HQ266579</td>
<td>β-Amyrin&lt;sup&gt;a&lt;/sup&gt;</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>149</td>
</tr>
<tr>
<td>BS</td>
<td>Vaccaria hispana</td>
<td>DQ915167</td>
<td>β-Amyrin</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>79</td>
</tr>
<tr>
<td>BPW</td>
<td>Betula platyphylla</td>
<td>AB055511</td>
<td>Lupeol</td>
<td>CCC</td>
<td>6/6/6/6/5</td>
<td>167</td>
</tr>
<tr>
<td>LUS</td>
<td>Bruguiera gymnorrhiza</td>
<td>AB289586</td>
<td>Lupeol</td>
<td>CCC</td>
<td>6/6/6/6/5</td>
<td>8</td>
</tr>
<tr>
<td>LUS1</td>
<td>Glycyrrhiza glabra</td>
<td>AB116228</td>
<td>Lupeol&lt;sup&gt;b&lt;/sup&gt;</td>
<td>CCC</td>
<td>6/6/6/6/5</td>
<td>45</td>
</tr>
<tr>
<td>KiLUS</td>
<td>Kalanchoe daigremontiana</td>
<td>HM623871</td>
<td>Lupeol&lt;sup&gt;c&lt;/sup&gt;</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>150</td>
</tr>
<tr>
<td>OSC3</td>
<td>Lotus japonicus</td>
<td>AB181245</td>
<td>Lupeol</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>120</td>
</tr>
<tr>
<td>OEW</td>
<td>Olea europea</td>
<td>AB025343</td>
<td>Lupeol&lt;sup&gt;d&lt;/sup&gt;</td>
<td>CCC</td>
<td>6/6/6/6/5</td>
<td>35</td>
</tr>
<tr>
<td>TRW</td>
<td>Taraxacum officinale</td>
<td>AB025345</td>
<td>Lupeol</td>
<td>CCC</td>
<td>6/6/6/6/5</td>
<td>136</td>
</tr>
<tr>
<td>MRN1</td>
<td>Arabidopsis thaliana</td>
<td>At5g42600</td>
<td>Marneral&lt;sup&gt;b&lt;/sup&gt;</td>
<td>CB</td>
<td>—/—/—/6</td>
<td>162</td>
</tr>
<tr>
<td>THAS1</td>
<td>Arabidopsis thaliana</td>
<td>At5g48010</td>
<td>Thalianol&lt;sup&gt;b&lt;/sup&gt;</td>
<td>CCC</td>
<td>6/6/5</td>
<td>27</td>
</tr>
<tr>
<td>SH1</td>
<td>Aster tataricus</td>
<td>AB609123</td>
<td>Shionone&lt;sup&gt;e&lt;/sup&gt;</td>
<td>CCC</td>
<td>6/6/6/6</td>
<td>121</td>
</tr>
<tr>
<td>KiGLS</td>
<td>Kalanchoe daigremontiana</td>
<td>HM623869</td>
<td>Glutinol&lt;sup&gt;f&lt;/sup&gt;</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>150</td>
</tr>
<tr>
<td>KiFRS</td>
<td>Kalanchoe daigremontiana</td>
<td>HM623870</td>
<td>Friedelin&lt;sup&gt;f&lt;/sup&gt;</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>150</td>
</tr>
<tr>
<td>KiTAS</td>
<td>Kalanchoe daigremontiana</td>
<td>HM623868</td>
<td>Taraxerol&lt;sup&gt;f&lt;/sup&gt;</td>
<td>CCC</td>
<td>6/6/6/6/6</td>
<td>150</td>
</tr>
<tr>
<td>IMS1</td>
<td>Luffa cylindrica</td>
<td>AB058643</td>
<td>Isomultiflorenol</td>
<td>CCC</td>
<td>6/6/6/6</td>
<td>43</td>
</tr>
<tr>
<td>OJAS</td>
<td>Oryza sativa</td>
<td>AK067451</td>
<td>Isoarborinol</td>
<td>CBC</td>
<td>6/6/6/5</td>
<td>164</td>
</tr>
<tr>
<td>PNA</td>
<td>Panax ginseng</td>
<td>AB265170</td>
<td>Dammaranediol II</td>
<td>CCC</td>
<td>6/6/6/5</td>
<td>144</td>
</tr>
<tr>
<td>BOS</td>
<td>Stevia rebaudiana</td>
<td>AB455264</td>
<td>Baccharis oxide</td>
<td>CCC</td>
<td>6/6/6/6</td>
<td>134</td>
</tr>
<tr>
<td>BARS1</td>
<td>Arabidopsis thaliana</td>
<td>At4g15370</td>
<td>Mixed products&lt;sup&gt;1&lt;/sup&gt;+22 other minor ones</td>
<td>CCC</td>
<td>6/6/6/6</td>
<td>77</td>
</tr>
<tr>
<td>CAMS1</td>
<td>Arabidopsis thaliana</td>
<td>At1g78955</td>
<td>Mixed products&lt;sup&gt;2&lt;/sup&gt;,&lt;sup&gt;3&lt;/sup&gt;,&lt;sup&gt;4&lt;/sup&gt;</td>
<td>CCC</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>LUP1</td>
<td>Arabidopsis thaliana</td>
<td>At1g78970</td>
<td>Mixed products&lt;sup&gt;5&lt;/sup&gt;,&lt;sup&gt;6&lt;/sup&gt;,&lt;sup&gt;8&lt;/sup&gt;,&lt;sup&gt;9&lt;/sup&gt;</td>
<td>CCC</td>
<td>6/6/6/6/5</td>
<td>46</td>
</tr>
<tr>
<td>LUP2</td>
<td>Arabidopsis thaliana</td>
<td>At1g78960</td>
<td>Mixed products&lt;sup&gt;4&lt;/sup&gt;,&lt;sup&gt;8&lt;/sup&gt;,&lt;sup&gt;7&lt;/sup&gt;,&lt;sup&gt;10&lt;/sup&gt;,&lt;sup&gt;11&lt;/sup&gt;,&lt;sup&gt;12&lt;/sup&gt;,&lt;sup&gt;13&lt;/sup&gt;,&lt;sup&gt;9&lt;/sup&gt;,&lt;sup&gt;14&lt;/sup&gt;</td>
<td>CCC</td>
<td>73</td>
<td></td>
</tr>
<tr>
<td>LUP5</td>
<td>Arabidopsis thaliana</td>
<td>At1g66960</td>
<td>Mixed products&lt;sup&gt;7&lt;/sup&gt;,&lt;sup&gt;UC&lt;/sup&gt;</td>
<td>CCC</td>
<td>6/6/6/5</td>
<td>25</td>
</tr>
<tr>
<td>PEN1</td>
<td>Arabidopsis thaliana</td>
<td>At4g15340</td>
<td>Mixed products&lt;sup&gt;15&lt;/sup&gt;,&lt;sup&gt;16&lt;/sup&gt;</td>
<td>CCC</td>
<td>6/6/5</td>
<td>66, 161</td>
</tr>
<tr>
<td>PEN3</td>
<td>Arabidopsis thaliana</td>
<td>At5g36150</td>
<td>Mixed products&lt;sup&gt;17&lt;/sup&gt;,&lt;sup&gt;12&lt;/sup&gt;,&lt;sup&gt;18&lt;/sup&gt;,&lt;sup&gt;19&lt;/sup&gt;,&lt;sup&gt;20&lt;/sup&gt;,&lt;sup&gt;21&lt;/sup&gt;</td>
<td>CCC</td>
<td>87</td>
<td></td>
</tr>
<tr>
<td>PEN6</td>
<td>Arabidopsis thaliana</td>
<td>At1g78500</td>
<td>Mixed products&lt;sup&gt;11&lt;/sup&gt;,&lt;sup&gt;10&lt;/sup&gt;,&lt;sup&gt;14&lt;/sup&gt;,&lt;sup&gt;UC&lt;/sup&gt;</td>
<td>CCC</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>CoOSC2</td>
<td>Castus spectus</td>
<td>AB058508</td>
<td>Mixed products&lt;sup&gt;10&lt;/sup&gt;,&lt;sup&gt;6&lt;/sup&gt;,&lt;sup&gt;8&lt;/sup&gt;,&lt;sup&gt;UC&lt;/sup&gt;</td>
<td>CCC</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>MS</td>
<td>Kandelia candel</td>
<td>AB257507</td>
<td>Mixed products&lt;sup&gt;10&lt;/sup&gt;,&lt;sup&gt;14&lt;/sup&gt;,&lt;sup&gt;4&lt;/sup&gt;</td>
<td>CCC</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>AMY2</td>
<td>Lotus japonicus</td>
<td>AF478455</td>
<td>Mixed products&lt;sup&gt;10&lt;/sup&gt;,&lt;sup&gt;4&lt;/sup&gt;,&lt;sup&gt;UC&lt;/sup&gt;</td>
<td>CCC</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>OSC8</td>
<td>Oryza sativa</td>
<td>AK070534</td>
<td>Mixed products&lt;sup&gt;22&lt;/sup&gt;,&lt;sup&gt;UC&lt;/sup&gt;</td>
<td>CCC</td>
<td>6/—/—/6</td>
<td>56</td>
</tr>
<tr>
<td>OSCPSM</td>
<td>Pisum sativum</td>
<td>AB034803</td>
<td>Mixed products&lt;sup&gt;4&lt;/sup&gt;,&lt;sup&gt;4&lt;/sup&gt;,&lt;sup&gt;2&lt;/sup&gt;,&lt;sup&gt;3&lt;/sup&gt;,&lt;sup&gt;9&lt;/sup&gt;,&lt;sup&gt;12&lt;/sup&gt;,&lt;sup&gt;10&lt;/sup&gt;,&lt;sup&gt;6&lt;/sup&gt;,&lt;sup&gt;8&lt;/sup&gt;</td>
<td>CCC</td>
<td>85</td>
<td></td>
</tr>
</tbody>
</table>


Table 1  (Continued)

<table>
<thead>
<tr>
<th>OSC</th>
<th>Species</th>
<th>GenBank ID/gene number</th>
<th>Product</th>
<th>Fold</th>
<th>Ring system</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RsM1</td>
<td>Rhizophora stylosa</td>
<td>AB263203</td>
<td>Mixed products</td>
<td></td>
<td>CCC</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6, 4, 10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RsM2</td>
<td>Rhizophora stylosa</td>
<td>AB263204</td>
<td>Mixed products</td>
<td></td>
<td>CCC</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6, 4, 10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTS2</td>
<td>Solanum lycopersicum</td>
<td>HQ266580</td>
<td>Mixed products</td>
<td></td>
<td>CCC</td>
<td>149</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4, 4, 23</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Superscript lowercase letters denote respective OSC transcript abundance in specific tissues as determined by reverse transcription polymerase chain reaction (RT-PCR), semiquantitative PCR, or northern blot analysis: aall tissues; broot tip (epidermis); cleaf and fruit epidermis; dnodules; eepicuticular layers of leaves and stems. Superscript numbers denote the products of OSCs that generate mixed products, with the numbers for each OSC listed in the order of the products’ abundance in gas chromatography–mass spectrometry (most abundant product listed first): 1baruol; 2camelliol C; 3achilleol A; 4β-amyrin; 5β,20-dihydroxylupane; 6germanicol; 7tirucalla-7,21-dien-3-β-ol; 8taraxasterol; 9Ψ-taraxasterol; 10lupeol; 11baueranol; 12butyrospermol; 13multiflorenol; 14α-amyrin; 15(3S,13R)-malabarica-17,21-dien-3,14-diol (arabidiol); 16(3S,21S)-malabarica-17-en-20,21-epoxy-3,14-diol (arabidiol 20,21-epoxide); 17tirucalla-7,24-dien-3-ol; 18tirucallol; 19 isostrictacol; 20(3S,13H)-malabarica-14(27),17,21-trien-3-ol; 21dammara-20,24-dien-3-ol; 22achilleol B; 23β-amyrin; 24taraxerol; additional uncharacterized products. Asterisks denote a predicted product. Abbreviations: CB, chair-boat; CBC, chair-boat-chair; CCC, chair-chair-chair.

Some other plant triterpene synthases produce different types of scaffolds through unusual reactions such as formation of a ketone at the C-3 position (96, 121), introduction of an oxide bridge (134), or ring cleavage to give seco-triterpenes (“seco” is derived from “sec,” the Latin for “to cut”) (24, 135, 162) (Figure 4a–c). Triterpene synthases can also be made to generate different products following incubation with unnatural substrates. For example, the A. thaliana OSC LUP1 is a multifunctional triterpene synthase that normally converts 2,3-oxidosqualene to lupeol, pentacyclic triterpenes, and a variety of triterpene alcohols and diols. Incubation of LUP1 with the alternative substrate 2,3-22,23-dioxidosqualene, generated by squalene monooxygenase–mediated synthesis using 2,3-oxidosqualene as a substrate, results in the formation of other diverse heterocyclic structures that are produced via the epoxy dammarenalen cation (126, 129) (Figure 4d).

Structural and Functional Analysis of Oxidosqualene Cyclases

Crystal structures for plant sterol- or triterpene-synthesizing OSCs are not yet available. However, the structures of two other relevant enzymes—SHC from the bacterium Alicyclobacillus acidocaldarius (110, 154, 155) and human lanosterol synthase (LAS) (145)—have been experimentally determined. Although these two enzymes share only ~25% amino acid identity, they have very similar architectures, both having two highly conserved (αα) barrel domains (known as the βγ fold) and a hydrophobic membrane-insertion helix (16, 98) (Figure 5). The substrate and cyclization product likely enter and leave the active site of the enzyme via the membrane (110, 145).

Investigations of SHC and LAS have led to the identification of residues implicated in the initiation of cyclization, ring formation, and potential stabilization of carbocationic intermediates (1, 125, 156, 158–160). For example, alanine-scanning mutagenesis of the aromatic residues in the hydrophobic cavity of SHC resulted in altered product profiles, thereby establishing the roles of different residues in product formation (84, 117). Site-saturation mutagenesis of the termination residue His234 in yeast ergosterol synthase (His232 in human LAS) also resulted in a range of cyclic products (159). This residue has been postulated to deprotonate the tetracyclic C-8/9 cation in lanosterol and cycloartenol synthases. Plant triterpene synthases such as lupeol and β-amyrin synthase have an aromatic residue at the corresponding position, which has been proposed to
Figure 3
Expression profiles and major products of *Arabidopsis thaliana* oxidosqualene cyclases (OSC). (a) Heat map showing expression profiles for 12 of the 13 *A. thaliana* OSC genes (transcripts for CAMSI were not detected, and so this gene was not included). Expression data were retrieved from Genevestigator V3 (48). (b) The major 2,3-oxidosqualene cyclization products made by each of the 13 *A. thaliana* OSCs. Some of these OSCs also make other products (Table 1).
stabilize the intermediary cation, thereby enabling further cyclization and ring expansion to pentacyclic triterpenes (109). In other studies, Matsuda and coworkers (76) combined two mutations (His477Asn and Ile481Val) to successfully convert the A. thaliana cycloartenol synthase CAS1 into an accurate lanosterol synthase, and Ebizuka and coworkers (72) used a combination of domain swapping and site-directed mutagenesis with lupeol and β-amyrin synthases to identify a motif (MWCYCR) within an 80-amino-acid region that determines β-amyrin formation. Two of these residues (Trp259 and Tyr261) are implicated in D/E-ring stabilization. These examples are not exhaustive, but they serve to illustrate the potential of combining sequence alignments and homology modeling with mutagenesis to probe the mechanisms of this fascinating family of enzymes. Investigations of the ways in which these enzymes generate such metabolic diversity from a single linear substrate will be a fertile area for future investigation, for both fundamental science and biotechnological applications. This would be greatly facilitated by experimental determination of the structures of exemplar plant sterol and triterpene synthases.

Phylogenetic Analysis of Functionally Characterized Plant Oxidosqualene Cyclases

Figure 6 shows a phylogenetic tree illustrating the relatedness of characterized plant OSCs. It is evident that the OSCs that generate the protosterol cation—the precursor for the sterols cycloartenol and lanosterol—all group together. These include OSCs from both dicots and monocots.

The triterpene synthases that generate the C-20 dammarenyl cation all group separately from the sterol synthases. These enzymes most likely arose directly or indirectly by duplication and divergence of cycloartenol synthase genes (28, 29, 53, 104, 107, 109, 164). It follows that the triterpene CCC fold and associated ability to cyclize 2,3-oxidosqualene to the dammarenyl cation are likely to be derived from the CBC fold of a sterol-producing progenitor, and that changes in substrate folding underlie this major functional diversification event. Four of the characterized triterpene synthases are from monocots: OsIAS, isoarborinol synthase from rice (Oryza sativa) (164); OsOSC8, also from rice, which makes achilleol B as its main cyclization product along with a variety of other triterpenes (56); β-amyrin synthase from diploid oat (Avena strigosa), which is required for the synthesis of the antimicrobial β-amyrin-derived triterpene glycosides known as avenacs (40, 107); and the multifunctional enzyme CsOSC2 from crape ginger (Costus speciosus), which generates β-amyrin, lupeol, germanicol, and additional uncharacterized cyclization products (61). These monocot triterpene synthases are more similar to the sterol synthases when compared with the other triterpene synthases shown in Figure 6, which are all from dicots, and form a discrete subgroup. It was previously reported that oat β-amyrin synthase (the only characterized β-amyrin synthase from monocots) evolved independently of the dicot β-amyrin synthases and shares greater similarity with the cycloartenol synthases (40, 107). This monocot β-amyrin synthase is clearly distinct from the dicot β-amyrin synthases, which form a large, distinct clade elsewhere in the tree (Figure 6).

Thus, the ability to cyclize 2,3-oxidosqualene to β-amyrin has arisen more than once in the plant kingdom. One other monocot OSC groups with the monocot triterpene synthase clade, namely O. sativa parkeol synthase (OsPS), which, like cucurbitadienol synthase from C. pepo, generates a sterol cyclization product (in this case parkeol) (56). This contrasts with the dicot OSC cucurbitadienol synthase from C. pepo (CpCPQ), which makes cucurbitadienol, an alternative rearrangement product of the protosteryl cation (130). CPQ groups with the cycloartenol and lanosterol synthases (Figure 6).

The remainder of the OSCs shown in Figure 6 consist of a discrete clade of lupeol synthases; a clade of A. thaliana OSCs that make particular major products, in some cases along with multiple minor products (AtBARS1, AtPEN1, AtTHAS1, AtPEN3, ArPEN6, and AtMRN1) (25, 27, 77, 288).
a  Introduction of ketone functionality

b  Introduction of an oxide bridge

c  Ring breaking

d  Generation of heterocyclic triterpenes

Thimmappa et al.
Figure 5
Crystal structures of human lanosterol synthase (LAS) (Protein Data Bank ID 1w6K) and bacterial squalene-hopene cyclase (SHC) (Protein Data Bank ID 2sqc). These enzymes have the βγ fold typical of class II terpene synthases (β and γ domains shown in green and yellow, respectively) and a hydrophobic membrane-insertion helix (shown in red) that is conserved across LAS, SHC, and other oxidosqualene cyclases.

87, 161, 162); and a broad group of OSCs, many of them multifunctional, that make a range of cyclization products (see Table 1 for references). Multifunctional OSCs may conceivably represent evolutionary intermediates (104) that are in the process of being refined to become more accurate pentacyclic triterpene synthases. As more characterized OSCs emerge, we will be able to draw on and augment this phylogenetic framework and thereby continue to develop our ability to predict product profiles based on DNA sequences.

Oxygenation of the Triterpene Scaffold
Although simple triterpenes such as β-amyrin and lupeol are common in plants, triterpene scaffolds are often further modified to more elaborate molecules by tailoring enzymes (6, 88, 99, 119). Cytochrome P450–mediated oxygenation of the scaffold (e.g., introduction of hydroxyl, ketone, aldehyde, carboxyl, or epoxy groups) is common. The functional groups introduced by such modifications may then pave the way for further tailoring by enzymes such as sugar transferases and acyltransferases. P450s therefore play a key role in functionalizing the triterpene scaffold.
Figure 6

Neighbor-joining tree of oxidosqualene cyclases (OSCs) from diverse plant species. OSC amino acid sequences (Table 1) were aligned using ClustalW with default parameters as implemented in the program MEGA5 (143). All positions containing gaps and missing data were eliminated. Evolutionary distances were computed using the JTT matrix-based method (59). Evolutionary analyses were conducted in MEGA5 with 1,000 bootstrap replicates (143). The scale bar (bottom, center) indicates 0.1 amino acid substitutions per site. The outgroup sequences used were cycloartenol synthases from *Chlamydomonas reinhardtii* (GenBank ID EDP09612) (80), *Dictyostelium discoideum* (dictyBase ID DDB_G0269226) (34), and the methanotropic bacterium *Stigmatella aurantiaca* (GenBank ID CAD39196) (9) and lanosterol synthases from *Homo sapiens* (GenBank ID AAC50184) (145), *Stigmatella aurantiaca* (GenBank ID EAU65610) (9), and the methanotropic bacterium *Methylococcus capsulatus* (GenBank ID AAU91075) (75). Cycloartenol synthases are shown in blue; monocot and dicot β-amyrin synthases are shown in red. Multifunctional OSCs are marked with an asterisk.
The plant P450s are a large and diverse group of enzymes that have been assigned to 11 clades based on sequence similarity. The P450s within these clades are further subdivided into families (95). Until recently, very little was known about the types of P450 enzymes involved in triterpene modifications. Within the past five years, however, there has been considerable activity in this area, and a total of 20 triterpene-modifying P450s have now been reported, 19 of them from dicots (18, 20, 28, 29–32, 37, 38, 49, 69, 70, 108, 127, 128, 131) (Table 2). Most of those characterized so far modify β-amyrin-based scaffolds (Figure 7a), although P450s that oxygenate other triterpene scaffolds—such as α-amyrin, lupeol, dihydro-lupeol, dammarenediol II, thalianol, arabiolid, marneral, and marnerol—have also been reported (Figure 7b). The enzymes shown in Table 2 belong to the CYP51, CYP71, CYP72, and CYP85 clans. The expansion of the latter three clans in plants is associated with metabolic diversification as plants colonized land (36, 82, 95). In contrast, members of the CYP51 clan are normally highly conserved and are associated with primary sterol biosynthesis. The limited number of P450s characterized to date makes it difficult to draw conclusions about the origins of triterpene-modifying P450s, but recruitment of P450s for triterpene biosynthesis is likely to have occurred multiple times during evolution.

The single example of a triterpene-modifying P450 from monocots [CYP51H10 from diploid oat (Avena strigosa)] belongs to the CYP51 clan (32, 108). The CYP51 clan is regarded as one of the most ancient of the P450 clans (95). Until recently, CYP51 enzymes were only known to have a single and highly conserved function—as sterol 14α-demethylases in the synthesis of essential sterols. The oat CYP51H10 enzyme is the first member of the CYP51 clan to have a different function—in the modification of the β-amyrin scaffold during the synthesis of antimicrobial triterpene glycosides known as avenacins. CYP51H10 belongs to a newly defined divergent group of CYP51 enzymes known as the CYP51H subfamily, which appears to be restricted to monocots and which also includes nine members of unknown function in rice (53, 108). The oat enzyme likely arose by duplication and neofunctionalization of a sterol 14α-demethylase gene (108). This enzyme is able to carry out two modifications to the β-amyrin scaffold (32, 70), specifically addition of a 16-hydroxyl group to the D ring and introduction of a 12-13 epoxide to the C ring (32). Molecular modeling and docking experiments suggest that C-16 hydroxylation precedes C-12,C-13 epoxidation (32). The 12-13 epoxide group is critical for the antimicrobial activity of avenacins (32). Multifunctional P450 enzymes from other P450 families that catalyze both hydroxylation and epoxidation reactions were previously known in microbes. To our knowledge, however, the oat CYP51H10 enzyme is the first plant P450 to catalyze both of these modifications. Given the close biogenic relationship between sterols and triterpenes, it may well be that other members of the divergent CYP51H subfamily (such as those in rice) are able to modify triterpene scaffolds. Functional analysis of this P450 subfamily in monocots is likely to be a fertile area for future investigation.

**Triterpene Glycosylation**

Triterpenes are often present in plants in glycosylated form. Glycosylation results in increased polarity and is often associated with bioactivity (6). Glycosylated triterpenes are also referred to as saponins. Many triterpenes have one or more sugar chains, normally attached at the C-3 and/or C-28 positions (although glycosylation at the C-4, C-16, C-20, C-21, C-22, and/or C-23 positions can also occur). These sugar chains are usually composed of glucose, galactose, arabinose, rhamnose, xylose, and glucuronic acid (although other sugars may also be incorporated) and are added onto hydroxyl groups or carboxyl groups, forming sugar acetals and sugar esters, respectively (147). Oligosaccharide chain formation is believed to occur through successive addition of sugars rather than transfer of a preformed sugar chain en bloc (6, 99, 119), and the
<table>
<thead>
<tr>
<th>Clan</th>
<th>Gene</th>
<th>Species</th>
<th>GenBank ID</th>
<th>Substrate</th>
<th>Reaction</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP51</td>
<td>CYP51H10</td>
<td>Avena strigosa</td>
<td>ABG88961</td>
<td>β-Amyrin</td>
<td>C-16 hydroxylation, C-12–C-13 epoxidation</td>
<td>32, 70, 108</td>
</tr>
<tr>
<td>CYP71</td>
<td>CYP71A16</td>
<td>Arabidopsis thaliana</td>
<td>NP_199073</td>
<td>Marneral, marnerol</td>
<td>C-23 hydroxylation</td>
<td>20, 28</td>
</tr>
<tr>
<td></td>
<td>CYP71D353</td>
<td>Lotus japonicus</td>
<td>KF460438</td>
<td>Dihydrolupeol, 20-hydroxylupeol</td>
<td>C-20 hydroxylation, C-28 oxidation (three-step oxidation)</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>CYP93E1</td>
<td>Glycine max</td>
<td>BAE94181</td>
<td>β-Amyrin, sophoradiol</td>
<td>C-24 hydroxylation</td>
<td>131</td>
</tr>
<tr>
<td></td>
<td>CYP93E2</td>
<td>Medicago truncatula</td>
<td>ABC59085</td>
<td>β-Amyrin</td>
<td>C-24 hydroxylation</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>CYP93E3</td>
<td>Glycyrrhiza uralensis</td>
<td>BAG68930</td>
<td>β-Amyrin</td>
<td>C-24 hydroxylation</td>
<td>127</td>
</tr>
<tr>
<td></td>
<td>CYP705A1</td>
<td>Arabidopsis thaliana</td>
<td>NP_195268</td>
<td>Arabidiol</td>
<td>C-15–C-16 cleavage</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>CYP705A5</td>
<td>Arabidopsis thaliana</td>
<td>Q9FI39</td>
<td>7β-Hydroxythalianol</td>
<td>C-15–C-16 desaturation</td>
<td>29</td>
</tr>
<tr>
<td>CYP72</td>
<td>CYP72A61v2</td>
<td>Medicago truncatula</td>
<td>BAL45199</td>
<td>24-Hydroxy-β-amyrin</td>
<td>C-22 hydroxylation</td>
<td>31, 128</td>
</tr>
<tr>
<td></td>
<td>CYP72A63</td>
<td>Medicago truncatula</td>
<td>BAL45200</td>
<td>β-Amyrin</td>
<td>C-30 hydroxylation, C-30 oxidation (three-step oxidation)</td>
<td>128</td>
</tr>
<tr>
<td></td>
<td>CYP72A68v2</td>
<td>Medicago truncatula</td>
<td>BAL45204</td>
<td>Oleanolic acid</td>
<td>C-23 hydroxylation (three-step oxidation)</td>
<td>31, 128</td>
</tr>
<tr>
<td></td>
<td>CYP72A134</td>
<td>Glycyrrhiza uralensis</td>
<td>BAL45207</td>
<td>β-Amyrin, 11-oxo-β-amyrin</td>
<td>C-30 hydroxylation (β-amyrin); C-30, C-22, and C-29 oxidation (11-oxo-β-amyrin)</td>
<td>128</td>
</tr>
<tr>
<td>CYP85</td>
<td>CYP88D6</td>
<td>Glycyrrhiza uralensis</td>
<td>BAG68929</td>
<td>β-Amyrin, 30-hydroxy-amyrin</td>
<td>C-11 oxidation (two-step oxidation)</td>
<td>127</td>
</tr>
<tr>
<td></td>
<td>CYP708A2</td>
<td>Arabidopsis thaliana</td>
<td>Q8L7D5</td>
<td>Thalianol</td>
<td>C-7 hydroxylation</td>
<td>20, 29</td>
</tr>
<tr>
<td></td>
<td>CYP716A12</td>
<td>Medicago truncatula</td>
<td>ABC59076</td>
<td>β-Amyrin, α-amyrin, lupeol</td>
<td>C-28 oxidation (three-step oxidation)</td>
<td>18, 30</td>
</tr>
<tr>
<td></td>
<td>CYP716A15</td>
<td>Vitis vinifera</td>
<td>BAJ84106</td>
<td>β-Amyrin, α-amyrin, lupeol</td>
<td>C-28 oxidation (three-step oxidation)</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>CYP716A17</td>
<td>Vitis vinifera</td>
<td>BAJ84107</td>
<td>β-Amyrin</td>
<td>C-28 oxidation (three-step oxidation)</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>CYP716A47</td>
<td>Panax ginseng</td>
<td>AYE75217</td>
<td>Dammarenediol II</td>
<td>C-12 hydroxylation</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>CYP716A53c2</td>
<td>Panax ginseng</td>
<td>AFO63031</td>
<td>Protopanaxadiol</td>
<td>C-6 hydroxylation</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>CYP716A1</td>
<td>Catharanthus roseus</td>
<td>AEX07773</td>
<td>β-Amyrin, α-amyrin, lupeol</td>
<td>C-28 oxidation (three-step oxidation)</td>
<td>49</td>
</tr>
</tbody>
</table>
recent demonstration of sequential glycosylation of steroidal alkaloids is consistent with this (54). The number of sugar chains, their composition, and their position on the triterpene scaffold provide considerable potential for metabolic diversification.

The enzymes that have so far been shown to glycosylate triterpenes belong to Carbohydrate-Active Enzymes (CAZy) family 1, as defined by Cantarel et al. (15). The family 1 glycosyltransferases are one of the largest groups of natural product–decorating enzymes in higher plants. The expansion of this family in higher plants reflects chemical diversification during the adaptation of plants to life on land (17, 165). These glycosyltransferases transfer sugars from nucleotide diphosphate-activated sugar moieties, usually UDP-glucose, to small hydrophobic acceptor molecules, and are referred to as family 1 UDP glycosyltransferases (UGTs) (10, 148). Other sugar donors include UDP-galactose, UDP-rhamnose, UDP-xylate, UDP-glucuronate, and UDP-arabinose. UGTs can be either highly selective or promiscuous in terms of the range of acceptors they recognize. The major principle governing acceptor recognition by UGTs seems to be regioslectivity (systematic glycosylation of the same position) rather than specificity for one or several structurally related compounds (19, 39, 148).

To date, 12 UGTs that are able to glycosylate triterpenes have been reported from various plant species (4, 5, 79, 94, 123, 131) (Table 3). Of these, 9 belong to the UGT73 family; the other 3 belong to the UGT71, UGT74, and UGT91 families, respectively. Figure 8 shows examples of the types of reactions catalyzed by these enzymes. These reactions include addition of glucose to the C-3 hydroxyl or the C-28 carboxylic acid groups of hederagenin and other oleanane scaffolds, as well as sugar-sugar additions such as transfer of a glucose onto a C-28 arabinose and additions of galactose and rhamnose to a C-3 sugar (4, 5, 123, 133) (Figure 8, Table 3). The majority of these enzymes are from the UGT73 family. It is noteworthy that other UGT73 enzymes use steroidal compounds as acceptors, including steroids and steroidal alkaloids (SaGT4A, SGT1, SGT2, and SGT3) and brassinosteroids (UGT73C5 and UGT73C6) (51, 54, 55, 64, 65, 83, 106, 139, 151, 152). The identification of further enzymes that are able to glycosylate triterpene scaffolds in different positions and build oligosaccharide chains will be important in manipulating the physicochemical and biological properties of these molecules.

Other Tailoring Enzymes

Triterpene scaffolds can undergo various other modifications in addition to oxygenation and glycosylation. For example, antimicrobial triterpene glycosides that are produced in oat roots (avenacins) are acylated at the C-21 position with either N-methyl anthranilate or benzoate (see Figure 7a for the structure of the major oat root triterpene,avenacin A-1). Mugford et al. (92) recently showed that the serine carboxypeptidase-like acyltransferase AsSCPL1 is responsible for this acylation step. In contrast to the BAHD acyltransferases, which utilize coenzyme A (CoA)–thioesters as the acyl donor, SCPL-like acyltransferases use O-glucose esters (81, 90, 91, 138). The sugar donors used by AsSCPL1 are N-methyl anthranilate-O-glucose and benzoyl-O-glucose (92). N-Methyl anthranilate is synthesized from anthranilate by the anthranilate N-methyltransferase AsMT1 (89). The family 1 glycosyltransferase (AsUGT74H5) then catalyzes the glycosylation of N-methyl anthranilate to N-methyl anthranilate-O-glucose, and the related enzyme AsUGT74H6 converts benzoate to benzoyl-O-glucose (102). Intriguingly, the three genes encoding AsSCPL1, AsMT1, and AsUGT74H5 are immediately adjacent to one another and form part of a larger biosynthetic cluster for averacin synthesis that also includes genes for the β-amyrin synthase AsBAS1 and the β-amyrin-modifying P450 AsCYP51H10 (32, 89, 92, 93, 102, 107, 108). C-21 acylation is a common feature of many of the most cytotoxic triterpene glycosides (105). Consistent with this, acylation is important for the potent biological activity of averacins (89). This set of
Figure 7
Modification of triterpene scaffolds by cytochrome P450s. (a) Characterized P450 enzymes that modify the β-amyrin scaffold. (b) Characterized P450 enzymes that modify other scaffolds. Multiple arrows indicate multiple sequential reactions catalyzed by the corresponding P450. The colors indicate the P450 clan to which each enzyme belongs: purple, CYP51; blue, CYP71; green, CYP72; red, CYP85. Table 2 provides more information about these enzymes.
Figure 7
(Continued)
Table 3  Triterpene glycosyltransferases (GTs)

<table>
<thead>
<tr>
<th>Family</th>
<th>GT</th>
<th>Species</th>
<th>GenBank ID</th>
<th>Substrate</th>
<th>Reaction</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>UGT71</td>
<td>UGT71G1</td>
<td><em>Medicago truncatula</em></td>
<td>AAW56092</td>
<td>Hederagenin</td>
<td>β-D-Glucosylation</td>
<td>4</td>
</tr>
<tr>
<td>UGT73</td>
<td>UGT73C10</td>
<td><em>Barbarea vulgaris</em></td>
<td>AFN26666</td>
<td>β-Amyrin, hederagenin</td>
<td>C-3 β-D-glucosylation</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>UGT73C11</td>
<td><em>Barbarea vulgaris</em></td>
<td>AFN26667</td>
<td>β-Amyrin, hederagenin</td>
<td>C-3 β-D-glucosylation</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>UGT73C12</td>
<td><em>Barbarea vulgaris</em></td>
<td>AFN26668</td>
<td>β-Amyrin, hederagenin</td>
<td>C-3 β-D-glucosylation</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>UGT73C13</td>
<td><em>Barbarea vulgaris</em></td>
<td>AFN26669</td>
<td>β-Amyrin, hederagenin</td>
<td>C-3 β-D-glucosylation</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>UGT73K1</td>
<td><em>Medicago truncatula</em></td>
<td>AAW56091</td>
<td>Hederagenin, soyasapogenol B/E</td>
<td>β-D-Glucosylation</td>
<td>4</td>
</tr>
<tr>
<td>UGT73F3</td>
<td>UGT73F3</td>
<td><em>Medicago truncatula</em></td>
<td>ACT34898</td>
<td>Hederagenin</td>
<td>C-28 β-D-glucosylation</td>
<td>94</td>
</tr>
<tr>
<td>UGT73F2</td>
<td>UGT73F2</td>
<td><em>Glycine max</em></td>
<td>BAM29362</td>
<td>Saponin A0-αg</td>
<td>C-3′ β-D-glucosylation</td>
<td>123</td>
</tr>
<tr>
<td>UGT73F4</td>
<td>UGT73F4</td>
<td><em>Glycine max</em></td>
<td>BAM29363</td>
<td>Saponin A0-αg</td>
<td>C-3′ β-D-xylosylation</td>
<td>123</td>
</tr>
<tr>
<td>UGT73P2</td>
<td>UGT73P2</td>
<td><em>Glycine max</em></td>
<td>BAJ99584</td>
<td>Soyasapogenol B 3-O-glucuronide</td>
<td>C-2′ β-D-galactosylation</td>
<td>133</td>
</tr>
<tr>
<td>UGT74</td>
<td>UGT74M1</td>
<td><em>Saponaria vaccaria</em></td>
<td>ABK76266</td>
<td>Gypsogenic acid</td>
<td>C-28 β-D-glucosylation</td>
<td>79</td>
</tr>
<tr>
<td>UGT91</td>
<td>UGT91H4</td>
<td><em>Glycine max</em></td>
<td>BAJ99585</td>
<td>Soyasaponin III</td>
<td>C-2′ β-D-rhamnosylation</td>
<td>133</td>
</tr>
</tbody>
</table>

enzymes for the synthesis of the acyl donor and subsequent transfer of the acyl group onto the triterpene scaffold is therefore likely to be a useful resource for triterpene modification.

**STRATEGIES FOR DISCOVERING NEW OXIDOSQUALENE CYCLASES AND TRITERPENE PATHWAYS**

**Oxidosqualene Cyclase Discovery**

Triterpenes are usually synthesized in particular plant tissues and/or at certain developmental stages. Production may also be induced in response to treatment with abiotic/biotic stresses or elicitors such as methyl jasmonate. For example, the triterpene glycoside saponins glycyrrhizin and avenacins accumulate in the roots of licorice (*Glycyrrhiza*) and oat (*Avena*), respectively, whereas in alfalfa (*Medicago sativa*), triterpene glycoside synthesis is induced in the leaves in response to pathogen and herbivore attack (6, 88, 99, 119). Treatment of cell suspension cultures with methyl jasmonate can provide a simple system for inducing and investigating triterpene biosynthesis in response to elicitor treatment (140). Most of the OSCs characterized to date have been cloned via expression-based strategies using degenerate primers and rapid amplification of cDNA ends (RACE)–polymerase chain reaction (PCR) or by screening cDNA libraries (6). Recent developments in high-throughput transcriptomics methods now open up powerful strategies for gene discovery that can be applied to diverse plant species.
Reactions catalyzed by characterized triterpene glycosyltransferases. UGT73C10 from *Barbarea vulgaris* and UGT73F3 from *Medicago truncatula* glucosylate the C-3 and C-28 positions of β-amyrin-derived (oleanane) triterpenes, respectively (5, 94). The *Saponaria vaccaria* enzyme UGT74M1 glucosylates the C-28 position of another oleanane triterpene, gypsogenic acid (79). Three glycosyltransferases that add sugars to triterpene glycosides are also shown. UGT73F2 from soybean (*Glycine max*) glucosylates the C-22-linked arabinose of soybean saponin A0-αg (123); UGT73P2, also from soybean, adds a galactose to the C-3-linked glucuronic acid of soyasapogenol-B, and a second soybean enzyme, UGT91H4, then adds a rhamnose to the galactose moiety (133).

The completion of the first plant genome sequence provided surprising insights into the potential of plants to make triterpenes. The *A. thaliana* genome contains 13 OSC genes, suggesting the presence of multiple OSC genes in plant genomes more generally. It has subsequently been established that the 13 *A. thaliana* OSCs all make different triterpenes (22, 25, 27, 46, 52, 66–68, 72, 73, 77, 87, 126, 132, 135, 141, 161, 162) (*Figure 3*). Genome mining has similarly revealed
the presence of 12 OSC genes in the rice genome (53). Of these, 1 encodes cycloartenol synthase (56); a further 7 are predicted to encode functional triterpene synthases (53), of which 3 have been functionally characterized and shown to make parkeol (OsPS), isoarborinol (OsIAS), and achilleol and a variety of other triterpenes (OsOSC8), respectively (56, 164). The functions of the other predicted rice triterpene synthases are unknown. From the genome sequences of these two species alone, it is clear that there is likely to be a huge and as yet unrealized capability for synthesis of diverse triterpene scaffolds hidden away in plant genomes. Further progress in unveiling this capability will inevitably be aided by the extension of new genomics technologies into more plant species and accessions.

**Oxidosqualene Cyclase Functional Analysis**

As indicated above, the primary routes for the discovery of new OSCs rely on expression-based approaches and increasingly, where genome sequence information is available, on genome mining (Figure 9). Once OSC sequences have been identified, they need to be functionally characterized. This is normally achieved through expression in yeast using strains that have been modified specifically for this purpose, either by engineering for elevated precursor supply or by using sterol auxotrophs that accumulate high levels of 2,3-oxidosqualene (6, 88, 99, 119). However, transient expression in leaves of the tobacco relative *Nicotiana benthamiana* using cowpea mosaic virus–based HyperTrans (CPMV-HT) technology has recently been shown to be an effective alternative means of expressing plant OSCs (32, 69). CPMV-HT has previously proved highly effective for the rapid, transient expression of a variety of structural proteins, including antibodies, vaccines, and empty viral particles (113–116, 122). In this system, the sequence to be expressed is inserted between a modified 5′ untranslated region and the 3′ untranslated region of CPMV RNA-2 in the T-DNA region of one of the series of *Agrobacterium tumefaciens* pEAQ vectors (116), which also encode the P19 suppressor of gene silencing. This method, which involves simply infiltrating *N. benthamiana* leaves with *A. tumefaciens* containing the appropriate expression constructs using a syringe without a needle, is very quick, and results can be obtained within six days. There is no requirement for engineering the earlier steps in the isoprenoid pathway in order to achieve expression of a heterologous OSC (oat β-amyrin synthase) in *N. benthamiana* leaves using this system, indicating that heterologous expression of this OSC can “pull through” the necessary precursors from generic host metabolism (32).

**Identification of Tailoring Enzymes**

Expression of OSCs in either yeast or *N. benthamiana* enables the corresponding triterpene scaffold(s) to be established using mass spectrometry and (where necessary) NMR spectrometry. These triterpene scaffolds may exist in unmodified form in plants (for example, in surface waxes) (11–13, 35, 58, 78, 142). Alternatively, they may be modified by tailoring enzymes. Expression-based approaches have enabled the discovery of genes encoding candidate tailoring enzymes (5, 18, 30, 31, 37, 38, 49, 79, 127, 128, 131). This is best achieved using the characterized OSC gene of interest as bait to search for other coexpressed genes with predicted functions in specialized metabolism (e.g., predicted P450 and glycosyltransferase genes). Functional validation of candidate tailoring enzymes is normally carried out in yeast.

An unexpected discovery has been that the genes for several specialized metabolic pathways, including triterpene pathways, are organized in clusters in plant genomes, a form of organization that is somewhat reminiscent of the natural product biosynthetic pathway gene clusters found in *Streptomyces* and some other bacteria (100). So far, such biosynthetic gene clusters have been reported
for diverse classes of plant natural products from different species, including both monocots and dicots (54, 63, 69, 157). The first of the plant triterpene biosynthetic clusters to be identified was the avenacin cluster from oat (103, 107), which contains a β-amyrin synthase gene, *AsbAS1*, along with genes for the tailoring enzymes CYP51H10 (a P450), AsSCPL1 (a serine carboxy-peptidase-like acyltransferase), AsMT1 (a methyltransferase), and AsUGT74H5 (a glucosyltransferase) (32, 40, 89, 92, 102, 107, 108). This cluster was first defined using a classical genetics approach that involved exploiting the unusual fluorescent properties of the major oat root avenacin to screen for
avenacin-deficient mutants of diploid oat (103), and was subsequently characterized by building a bacterial artificial chromosome contig spanning AsbAS1 (89, 92, 102, 107, 108). Other uncharacterized genes that are required for avanacyn synthesis have been defined by mutation and are also linked (93, 103). The avenacin cluster spans more than 200 kb. Osbourn and coworkers (28, 29) subsequently discovered two other triterpene biosynthetic clusters in A. thaliana by mining the genome for candidate biosynthetic gene clusters that contained OSC genes, and predicted and functionally validated the thalianol and marneral clusters (35 and 38 kb, respectively). An OSC gene and two linked, functionally associated P450 genes have also recently been identified in Lotus japonicus (69).

How common such clusters are in plant genomes remains to be established. Other new triterpene biosynthetic gene clusters will likely emerge as we learn more about the genomic context of OSC genes. For example, a candidate cluster consisting of an OSC gene, an acyltransferase gene, and two P450 genes, all of which are immediately adjacent to one another and coexpressed, is present in the cucumber genome. This cluster may be required for cucurbitacin synthesis (50).

The genes within these biosynthetic clusters are more closely related to plant genes than to microbial ones, and they are unlikely to have been introduced by horizontal gene transfer from microbes (63). Indeed, the evidence suggests that these clusters form in dynamic regions of chromosomes, presumably in response to strong selection (28). Clustering enables coinheritance of beneficial combinations of alleles that together confer a selective advantage (e.g., the production of defense compounds), and by enabling the pathway to be inherited intact, it may also ensure that bioactive/deleterious pathway intermediates are not unleashed. It further provides for regulation at the chromatin level, for which there is a growing body of evidence. We speculate that gene clustering is a feature of specialized metabolic pathways that have evolved relatively recently in evolutionary time. Comparative genomics is now beginning to provide the first insights into the dynamics of cluster formation (28).

Thus, the genes for triterpene biosynthetic pathways are coexpressed in particular plant parts, at certain growth stages, and/or in response to environmental triggers. Gene expression analysis provides a highly effective means of identifying genes for new OSCs as well as coexpressed genes for candidate tailoring enzymes. This strategy is independent of any knowledge of genomic context. However, if genome sequence is available, either from genome sequencing projects or from bacterial artificial chromosome contigs spanning OSC genes of interest, then positional information may prove to be extremely valuable for identifying gene clusters for entire triterpene biosynthetic pathways. Functional validation of new triterpene biosynthetic pathways can be demonstrated by coexpression of the OSC and candidate tailoring enzymes in yeast and/or N. benthamiana. The latter has enabled the expression of all five genes in the avenacin cluster and so has undergone proof of concept for expression of OSCs, P450s, acyltransferases, methyltransferases, and glycosyltransferases (32, 89, 92). There are some caveats associated with determining the biochemical functions of enzymes based on expression in heterologous hosts because this may give misleading results. Where possible, tests of function in the plant of origin using a combination of gene knockouts, RNA interference, and/or overexpression provide a complementary means of establishing the roles of enzymes, identifying pathway intermediates, and evaluating the likely biological/ecological significance of triterpene biosynthetic pathways (e.g., 18, 28, 29, 32, 40, 69, 89, 92, 93, 103, 107, 108, 123).

CONCLUSIONS

Until recently, the study of triterpenes has been hindered by the intractability of these molecules to chemical synthesis and by the dearth of enzymes available for bioengineering. Recent advances in our understanding of the genes and enzymes required for triterpene synthesis are now opening
up the potential to rapidly access these biologically and commercially important compounds. Few hypotheses currently exist regarding links between structure and activity for these diverse and complex molecules. To truly realize the scope of the metabolic diversity represented within the triterpene family, systematic studies will be needed. The development of platforms for synthetic biology–based generation and functionalization of triterpene scaffolds will enable the production of known and novel triterpenes and systematic comparison of the biological and chemical properties of these molecules.

Some of the most important insights into triterpene structure and function are likely to come from investigations of the roles of these compounds in their plants of origin. So far, investigations of this kind have been limited. Knowledge of the expression profiles of different types of OSCs will guide investigations of biochemical and biological function (Figure 3, Table 1). Simple triterpenes are known to accumulate in the epicuticular and intracuticular wax layers of stem and leaf surfaces and may have physiological roles in protection against dehydration and possibly against herbivores (11–13, 35, 58, 78, 142, 166). Lupeol synthases are also expressed preferentially in root nodules of certain legumes (23, 41, 57), and lupeol is implicated in the regulation of nodule development (23). The avenacin triterpene glycosides protect oat roots against attack by soil-borne fungal pathogens such as “take-all,” which causes major yield losses on wheat (103), and other β-amyrin-derived triterpene glycosides confer resistance to flea beetle (Phyllotreta nemorum) in Barbarea vulgaris, a member of the Brassicaceae family (74). Evidence is also emerging to suggest that triterpenes may have roles in plant growth and development (28, 29, 33). The discovery of regulators of triterpene biosynthesis will represent a further important area of future research because pathway-specific transcription factors have not yet been reported.

**SUMMARY POINTS**

1. Triterpenes are one of the most numerous and diverse groups of plant natural products and have a wide range of applications in the food, cosmetics, pharmaceutical, and industrial biotechnology sectors.

2. Triterpenes are complex molecules that are for the most part beyond the reach of chemical synthesis.

3. Triterpenes are synthesized from 2,3-oxidosqualene by oxidosqualene cyclases (OSCs), which collectively generate more than 100 different triterpene scaffolds; these scaffolds are then further modified by tailoring enzymes (e.g., cytochrome P450s, sugar transferases, and acyltransferases) to give even greater metabolic diversity.

4. The genes for triterpene biosynthetic pathways are—in at least some cases—organized in plant genomes as coexpressed metabolic gene clusters.

5. Major advances in our understanding of the genes, enzymes, and pathways required for the synthesis of these molecules are now opening up unprecedented opportunities for triterpene metabolic engineering.

6. Suitable heterologous systems for (re)construction of triterpene biosynthetic pathways are expression in yeast and Agrobacterium-mediated transient expression in Nicotiana benthamiana leaves.

7. The development of platforms for synthetic biology–based generation and functionalization of triterpene scaffolds will enable the generation of triterpene libraries and allow systematic comparison of the biological and chemical properties of these molecules.
DISCLOSURE STATEMENT

A.O. is a coinventor on patents filed on the five cloned avenacin genes. The authors are not aware of any other affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

A.O.’s laboratory is funded by awards from the Biotechnology and Biological Sciences Research Council (BBSRC) (BB/K005952/1), the Engineering and Physical Sciences Research Council (EP/K034359/1), and the European Union (KBBE-2013-7; TriForC); a BBSRC Institute Strategic Program Grant (“Understanding and Exploiting Plant and Microbial Metabolism,” BB/J004561/1); and the John Innes Foundation. P.O.’s laboratory is funded by an award from the BBSRC (BB/K003690/1) and BBSRC Institute Strategic Program Grants (BB/J004561/1 and “Food and Health,” BB/I015345/1). R.T. is funded by a European Commission Marie Curie International Incoming Fellowship (IF-275689: TRICYCLE).

LITERATURE CITED


155. Thimmappa et al.


Contents

From the Concept of Totipotency to Biofortified Cereals
Ingo Potrykus ................................................................. 1

The Structure of Photosystem II and the Mechanism of Water
Oxidation in Photosynthesis
Jian-Ren Shen ................................................................. 23

The Plastid Terminal Oxidase: Its Elusive Function Points to Multiple
Contributions to Plastid Physiology
Wojciech J. Nawrocki, Nicolas J. Tourasse, Antoine Taly, Fabrice Rappaport,
and Francis-André Wollman ................................................ 49

Protein Maturation and Proteolysis in Plant Plastids, Mitochondria,
and Peroxisomes
Klaas J. van Wijk ............................................................. 75

United in Diversity: Mechanosensitive Ion Channels in Plants
Eric S. Hamilton, Angela M. Schlegel, and Elizabeth S. Haswell .................. 113

The Evolution of Plant Secretory Structures and Emergence of
Terpenoid Chemical Diversity
Bernd Markus Lange .......................................................... 139

Strigolactones, a Novel Carotenoid-Derived Plant Hormone
Salim Al-Babili and Harro J. Bouwmeester ................................ 161

Moving Toward a Comprehensive Map of Central Plant Metabolism
Ronan Sulpice and Peter C. McKeown ................................... 187

Engineering Plastid Genomes: Methods, Tools, and Applications in
Basic Research and Biotechnology
Ralph Bock ........................................................................ 211

RNA-Directed DNA Methylation: The Evolution of a Complex
Epigenetic Pathway in Flowering Plants
Marjori A. Matzke, Tatsuo Kanno, and Antonius J.M. Matzke ..................... 243

The Polycomb Group Protein Regulatory Network
Iva Mozgova and Lars Hennig ................................................ 269
The Molecular Biology of Meiosis in Plants
Raphaël Mercier, Christine Mézard, Eric Jenczewski, Nicolas Macaisne, and Mathilde Grelon .......................................................... 297

Genome Evolution in Maize: From Genomes Back to Genes
James C. Schnable ........................................................................ 329

Oxygen Sensing and Signaling
Joost T. van Dongen and Francesco Licausi ................................ 345

Diverse Stomatal Signaling and the Signal Integration Mechanism
Yoshiyuki Murata, Izumi C. Mori, and Shintaro Munemasa ............. 369

The Mechanism and Key Molecules Involved in Pollen Tube Guidance
Tetsuya Higashiyama and Hidenori Takeuchi ................................ 393

Signaling to Actin Stochastic Dynamics
Jiejie Li, Laurent Blanchon, and Christopher J. Staiger .................. 415

Photoperiodic Flowering: Time Measurement Mechanisms in Leaves
Young Hun Song, Jae Sung Shim, Hannah A. Kinmonth-Schultz, and Takato Imaizumi ................................................................. 441

Brachypodium distachyon and Setaria viridis: Model Genetic Systems for the Grasses
Thomas P. Brutnell, Jeffrey L. Bennetzen, and John P. Vogel ............ 465

Effector-Triggered Immunity: From Pathogen Perception to Robust Defense
Haitao Cui, Kenichi Tsuda, and Jane E. Parker ............................... 487

Fungal Effectors and Plant Susceptibility
Libera Lo Presti, Daniel Lanver, Gabriel Schweizer, Shigeyuki Tanaka, Liang Liang, Marie Tollot, Alga Zuccaro, Stefanie Reissmann, and Regine Kahmann .......................................................... 513

Responses of Temperate Forest Productivity to Insect and Pathogen Disturbances
Charles E. Flower and Miguel A. Gonzalez-Meler .......................... 547

Plant Adaptation to Acid Soils: The Molecular Basis for Crop Aluminum Resistance
Leon V. Kochian, Miguel A. Pineros, Jiping Liu, and Jurandir V. Magalhaes .......................................................... 571

Terrestrial Ecosystems in a Changing Environment: A Dominant Role for Water
Carl J. Bernacchi and Andy VanLoocke ........................................ 599